

Subcritical multiplicative chaos for regularized counting statistics from random matrix theory

Article (Published Version)

Lambert, Gaultier, Ostrovsky, Dmitry and Simm, Nick (2018) Subcritical multiplicative chaos for regularized counting statistics from random matrix theory. *Communications in Mathematical Physics*, 360 (1). pp. 1-54. ISSN 0010-3616

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/73392/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



Subcritical Multiplicative Chaos for Regularized Counting Statistics from Random Matrix Theory

Gaultier Lambert¹, Dmitry Ostrovsky², Nick Simm³

¹ Department of Mathematics, Universität Zürich, Zurich, Switzerland. E-mail: gaultier.lambert@math.uzh.ch

² Stamford, CT, USA. E-mail: dm_ostrov@aya.yale.edu

³ Department of Mathematics, School of Mathematical and Physical Sciences, University of Sussex, Falmer Campus, Brighton BN1 9QH, UK. E-mail: N.J.Simm@sussex.ac.uk

Received: 3 January 2017 / Accepted: 30 November 2017

Published online: 16 April 2018 – © The Author(s) 2018

Abstract: For an $N \times N$ Haar distributed random unitary matrix U_N , we consider the random field defined by counting the number of eigenvalues of U_N in a mesoscopic arc centered at the point u on the unit circle. We prove that after regularizing at a small scale $\epsilon_N > 0$, the renormalized exponential of this field converges as $N \rightarrow \infty$ to a Gaussian multiplicative chaos measure in the whole subcritical phase. We discuss implications of this result for obtaining a lower bound on the maximum of the field. We also show that the moments of the total mass converge to a Selberg-like integral and by taking a further limit as the size of the arc diverges, we establish part of the conjectures in Ostrovsky (Nonlinearity 29(2):426–464, 2016). By an analogous construction, we prove that the multiplicative chaos measure coming from the sine process has the same distribution, which strongly suggests that this limiting object should be universal. Our approach to the L^1 -phase is based on a generalization of the construction in Berestycki (Electron Commun Probab 22(27):12, 2017) to random fields which are only *asymptotically* Gaussian. In particular, our method could have applications to other random fields coming from either random matrix theory or a different context.

1. Introduction

1.1. Background and results for the CUE. The study of Gaussian fields with logarithmic correlations has seen many recent developments in the last few years. One of those concerns a relation to the eigenvalues of random matrices, which can be traced back to a work of Hughes, Keating and O’Connell [37]. They studied the characteristic polynomial of random $N \times N$ matrices U_N sampled from the unitary group according to the Haar measure, also known as the Circular Unitary Ensemble (CUE). One of their key results

G. L. gratefully acknowledges the support of the grant KAW 2010.0063 from the Knut and Alice Wallenberg Foundation while at KTH, Royal Institute of Technology.

N. S. gratefully acknowledges financial support from a Leverhulme Trust Early Career Fellowship ECF-2014-309.

was to prove that $Z_N(\theta) = \sqrt{2} \log |\det(e^{i\theta} - U_N)|$ converges in law as $N \rightarrow \infty$ to a log-correlated Gaussian field which corresponds to the restriction of the two-dimensional Gaussian Free Field on the unit circle. This field can be represented as the Fourier series:

$$Z(\theta) = \frac{1}{\sqrt{2}} \operatorname{Re} \left\{ \sum_{k=1}^{\infty} \frac{e^{-ik\theta}}{\sqrt{k}} Z_k \right\}, \quad (1.1)$$

where Z_k are i.i.d. standard complex Gaussian variables. In particular, almost surely, this series does not converge in $L^2(\mathbb{T})$, and $Z(\theta)$ only exists as a random distribution, or a generalized random function with correlation kernel $\mathbb{E}(Z(\theta)Z(\theta')) = -\log |e^{i\theta} - e^{i\theta'}|$.

More recent developments concern the extreme value statistics of the field $Z_N(\theta)$ and on the related issue of making sense of its exponential in the limit as $N \rightarrow \infty$. The authors of [26, 27] gave a very precise conjecture for the maximum value of $Z_N(\theta)$, which has recently seen significant progress [1, 13, 57]. This conjecture is intimately related to the following¹ concerning the total mass of the field $Z_N(\theta)$.

Conjecture 1.1 (Fyodorov and Keating [27]). *Let $\gamma > 0$ and define*

$$M_N^\gamma := \int_0^{2\pi} e^{\gamma Z_N(\theta)} d\theta \quad (1.2)$$

For any $q \in \mathbb{N}$ such that $\gamma^2 q < 2$, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} N^{-\frac{\gamma^2 q}{2}} \mathbb{E}[(M_N^\gamma)^q] &= C_{\gamma, q} \int_{[0, 2\pi]^q} \prod_{1 \leq j < k \leq q} |e^{i\theta_j} - e^{i\theta_k}|^{-\gamma^2} d\theta_1 \dots d\theta_q \\ &= C_{\gamma, q} \frac{\Gamma(1 - \gamma^2 q/2)}{\Gamma(1 - \gamma^2/2)^q} \end{aligned} \quad (1.3)$$

where $C_{\gamma, q} = (2\pi)^q \frac{G(1+\gamma/\sqrt{2})^{2q}}{G(1+\gamma\sqrt{2})^q}$ and $G(z)$ is the Barnes G -function.

The case $q = 2$ of Conjecture 1.1 was solved by Claeys and Krasovsky [14] by a rather delicate asymptotic analysis of Toeplitz determinants with merging Fisher–Hartwig singularities.² For $q \neq 1, 2$, this conjecture remains an open problem. On the other hand, the normalization by $N^{-\frac{\gamma^2 q}{2}}$ on the left-hand side of (1.3) suggests that $N^{-\gamma^2/2} M_N^\gamma$ might converge in distribution to a non-trivial limiting random variable as $N \rightarrow \infty$, at least for some range of γ values. This has been shown by Webb [70] for any $\gamma < 1$. The condition $\gamma < 1$ is called the L^2 -phase because it is precisely the regime where the second moment (case $q = 2$ of (1.3)) is finite. Webb showed that the limiting random variable can be described in terms of *Gaussian multiplicative chaos* (GMC), a theory devoted to properly defining the exponential of a log-correlated Gaussian field. These exponentials are interpreted as random measures and are naturally linked to the geometric properties of the fields. This topic is currently under intense

¹ For convenience we state Conjecture 1.1 for $q \in \mathbb{N}$. The full conjecture given by these authors involves an analytic continuation to $q \in \mathbb{C}$ which allowed them to calculate the extreme value statistics of $Z_N(\theta)$ by invoking the so-called ‘freezing scenario’ expected of log-correlated Gaussian processes.

² Claeys and Krasovsky also obtain the asymptotics for the so-called critical and super-critical regimes of the second moment ($\gamma = 1$ and $\gamma > 1$ respectively). For general q with $\gamma^2 q > 2$ the asymptotic order of magnitude was conjectured by Fyodorov and Keating [27].

investigation, partly because of its applications in subjects such as Liouville quantum gravity, turbulence or mathematical finance, see [60] for a review. We provide a detailed discussion of GMC in Appendix B, but the basic idea can be summarised as follows. Consider a Gaussian field $G(u)$ on a domain $\mathfrak{A} \subset \mathbb{R}^d$ with mean zero and covariance

$$\mathbb{E}(G(u)G(v)) = \log \frac{1}{|u - v|} + g(u, v) \quad (1.4)$$

where $g(u, v)$ is bounded and continuous. Because of the logarithmic singularity in (1.4), $G(u)$ is not an ordinary function, so one wishes to regularize it in some way, obtaining $G_\epsilon(u)$ which is a smooth Gaussian field for every finite $\epsilon > 0$. Then one may consider the exponential of the field $G_\epsilon(u)$ and remove the diverging part in order to take the limit as $\epsilon \rightarrow 0$. Typically, one finds that this limit does not depend on the regularization procedure. This was first established by Kahane [41, 42], using martingale methods and more recently extended to other regularizations, such as convolutions.

Theorem 1.2 (Robert and Vargas [62], Berestycki [5]). *Let G be a log-correlated Gaussian field and ϕ a smooth probability density function (mollifier) on \mathbb{R}^d . For any $\epsilon > 0$, let $G_\epsilon = G * \phi_\epsilon$ be the convolution of G with $\phi_\epsilon(x) := \frac{1}{\epsilon} \phi(\frac{x}{\epsilon})$. For any function $w \in L^1(\mathbb{R}^d)$ uniformly bounded with compact support, define the family of random variables*

$$v_\epsilon^\gamma(w) = \int e^{\gamma G_\epsilon(u) - \frac{\gamma^2}{2} \text{Var}(G_\epsilon(u))} w(u) du. \quad (1.5)$$

Then for every $\gamma < \sqrt{2d}$, $v_\epsilon^\gamma(w)$ converges in $L^1(\mathbb{P})$ to a non-trivial random variable $v^\gamma(w)$ as $\epsilon \rightarrow 0$. This limit does not depend on the mollifier ϕ and v^γ defines a random measure which is called the GMC measure associated with the field G .

The aim of this paper is to establish proofs of this convergence for fields which are only *asymptotically* Gaussian and log-correlated. Our main interest in such fields will be those which naturally arise in random matrix theory, but our theory is likely to apply in other situations.

We now define our main objects of study for the CUE. Let $e^{i\theta_1}, \dots, e^{i\theta_N}$ denote the eigenvalues of a Haar distributed random unitary matrix U_N , with the convention that $\theta_1, \dots, \theta_N \in [-\pi, \pi)$. We consider the random process $u \rightarrow W_N(u)$ which counts the number of eigenangles in an interval centered around u :

$$W_N(u) = \sum_{j=1}^N \chi_u(N^\alpha \theta_j), \quad (1.6)$$

where $\chi_u(x) = \pi \mathbb{1}_{|x-u| \leq \ell/2}$ for some fixed $\ell > 0$. We emphasize that throughout the paper, the role played by u is that of the spatial variable of the random process under consideration. The parameter α is called the spectral scale and the process (1.6) can be studied in three different regimes. In the microscopic regime $\alpha = 1$, the field $W_N(u)$ converges weakly as $N \rightarrow \infty$ to a counting function for the sine process which is not a Gaussian process, but we will come back to this in Sect. 1.3. On the other hand, either in the mesoscopic regime³ $0 < \alpha < 1$ or in the global regime $\alpha = 0$, if centered, the process $W_N(u)$ will converge in a certain sense to a log-correlated Gaussian field. We

³ Mesoscopic means that the number of objects (zeroes, eigenvalues) that are being counted goes to infinity, whereas the length of the interval, over which they are being counted, goes to zero.

will focus on the mesoscopic regime where the geometry of the spectrum is unimportant, although our theory also applies for $\alpha = 0$ with minor changes. Inspired by Theorem 1.2, we will work with the following family of regularizations instead of working directly with the counting statistics (1.6). Let ϕ be a mollifier and define

$$X_{N,\epsilon}(u) := \sum_{j=1}^N (\chi_u * \phi_\epsilon)(N^\alpha \theta_j) \quad (1.7)$$

where $\phi_\epsilon(\theta) = \frac{1}{\epsilon} \phi\left(\frac{\theta}{\epsilon}\right)$. Throughout the paper, we use the notation $\overline{X}_{N,\epsilon_N} := X_{N,\epsilon_N} - \mathbb{E}X_{N,\epsilon_N}$ to denote a recentering by the expectation. The smoothed fields (1.7) were introduced in [55] in the context of counting statistics of the Riemann zeros. The parameter $\epsilon > 0$ controls the scale of regularization for $X_N(u)$, which we will allow to converge to zero as the dimension $N \rightarrow \infty$. Roughly speaking, the speed at which the sequence $\epsilon_N \rightarrow 0$ controls how close the centered field $\overline{X}_{N,\epsilon_N}$ is to a Gaussian process for large N .

For fixed $\epsilon > 0$, (1.7) is a smooth linear statistic and by Soshnikov's central limit theorem (CLT) [68], for every $u \in \mathbb{R}$, we have the convergence in distribution to a Gaussian random variable,

$$\overline{X}_{N,\epsilon}(u) \Rightarrow \mathcal{N}\left(0, \int_{\mathbb{R}} |k| \widehat{|\chi_u * \phi_\epsilon(k)|^2} dk\right), \quad (1.8)$$

as $N \rightarrow \infty$, see formula A.1 for our normalization of the Fourier transform. A direct calculation with the limiting variance in (1.8) easily shows (using *e.g.* the fact that the Fourier transform of χ_u decays as $1/k$ for large k) that $\text{Var}(X_{N,\epsilon}(u)) \sim \log \frac{1}{\epsilon}$ as $N \rightarrow \infty$ followed by $\epsilon \rightarrow 0$. In general, the covariance structure coming from Soshnikov's CLT is associated with an $H^{1/2}$ Sobolev space in the following way: for any functions $g, h \in C_0^1(\mathbb{R})$,

$$\begin{aligned} \lim_{N \rightarrow \infty} \text{Cov}\left(\sum_{j=1}^N g(N^\alpha \theta_j), \sum_{j=1}^N h(N^\alpha \theta_j)\right) &= \int_{\mathbb{R}} |k| \hat{g}(k) \hat{h}(-k) dk \\ &= \frac{1}{2\pi^2} \int_{\mathbb{R}^2} g'(x) h'(y) \log \frac{1}{|x - y|} dx dy. \end{aligned} \quad (1.9)$$

This is suggestive of a logarithmic covariance structure underlying the mesoscopic statistics of CUE eigenvalues. Applied to the statistic (1.7), one can easily show that for $u, v \in \mathbb{R}$ fixed ($u \neq v$), we have

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \text{Cov}(X_{N,\epsilon}(u), X_{N,\epsilon}(v)) = \log \frac{1}{|u - v|} + \log |\ell^2 - (u - v)^2|^{1/2}. \quad (1.10)$$

By analogy with (1.5), we are going to study the convergence of the following quantity as $N \rightarrow \infty$, where ϵ may converge to zero with N .

$$\mu_{N,\epsilon}^{\gamma, \text{CUE}}(w) := \int_{\mathbb{R}} e^{\gamma \overline{X}_{N,\epsilon}(u) - \frac{\gamma^2}{2} \text{Var}(\overline{X}_{N,\epsilon}(u))} w(u) du, \quad (1.11)$$

where $w \in L^1(\mathbb{R})$ has compact support.

Theorem 1.3. *Consider the regularized statistic (1.7) with $\ell > 0$ and mesoscopic scale $0 < \alpha < 1$ fixed. Suppose that $\epsilon_N \rightarrow 0$ in such a way that for some $\kappa > 0$,*

$$N^{\alpha-1} \epsilon_N^{-1} = O_{N \rightarrow \infty}(N^{-\kappa}). \quad (1.12)$$

Then for every $0 < \gamma < \sqrt{2}$ and $w \in L^1(\mathbb{R})$ uniformly bounded with compact support, we have the convergence in distribution

$$\mu_{N, \epsilon_N}^{\gamma, \text{CUE}}(w) \Rightarrow v^\gamma(w), \quad N \rightarrow \infty, \quad (1.13)$$

where v^γ is the GMC measure defined in Theorem 1.2 with $g(u, v) = \log |\ell^2 - |u - v|^2|^{1/2}$ in the covariance formula (1.4).

Remark 1.4. The reader may observe that $g(u, v)$ in Theorem 1.3 is not bounded below, as in the hypothesis of Theorem 1.2. However, it is easy to see that precisely the same steps carried out in [5] for constructing the multiplicative chaos measure show that this assumption is not required, the main crucial point being that $g(u, v)$ is bounded above.

Remark 1.5. In the following, we interpret v_ϵ^γ in (1.5) and $\mu_{N, \epsilon}^{\gamma, \text{CUE}}$ in (1.11) as absolutely continuous random measures and we will use the notation $v_\epsilon^\gamma(S) = v_\epsilon^\gamma(\mathbb{1}_S)$ for any compact Borel subset $S \subset \mathbb{R}^d$ and similarly for $\mu_{N, \epsilon_N}^{\gamma, \text{CUE}}$.

The condition (1.12) is rather natural and we will discuss its meaning and necessity in the next subsection, but we mention that its importance was first emphasized for smoothed statistics of the Riemann zeros in work of the second author [55]. There it is shown that (1.12) is the natural slow decay condition both for keeping the statistic mesoscopic and for preserving its $H^{1/2}$ -Gaussianity under smoothing.

It turns out that the appearance of the $H^{1/2}$ -Gaussian noise is remarkably universal in the mesoscopic limit of one dimensional ensembles with random matrix type repulsion and this problem has attracted renewed interest in the last couple of years. For instance, the analogue of Soshnikov's CLT (1.9) was obtained for the GUE [28], more general invariant ensembles [12, 46], Wigner matrices [23, 34, 47], β -ensembles [2, 10] and for zeros of the Riemann zeta function [11, 63]. It is likely the counterpart of Theorem 1.3 continues to hold for these models as well. To add some weight to this assertion, we will present the proof of Theorem 1.3 under some general criteria (see Theorem 1.7 below) and provide an analogous result for the sine process, that is the random point process which describes the microscopic limit of a wide class of Hermitian random matrices, see Sect. 1.3. Thus, from the point of view of Gaussian multiplicative chaos, one may view the different ensembles of random matrices as alternative ways of regularizing a log-correlated Gaussian field.

We also prove a result establishing the convergence of the q -moments of (1.11) and a relation to Selberg integrals. Although ℓ is fixed in Theorem 1.3, we now consider the case where $\ell = L(N) \rightarrow \infty$ as $N \rightarrow \infty$. Similarly, the rate at which $L(N) \rightarrow \infty$ is naturally restricted by the mesoscopic scale,

$$L(N) \log(N) = o_{N \rightarrow \infty}(N^\alpha). \quad (1.14)$$

Under this restriction we prove that the moments of the total mass are given by ratios of Gamma functions in the limit $N \rightarrow \infty$.

Theorem 1.6 (Moments of the total mass and Selberg integrals). *Let $0 < \alpha < 1$ and suppose that the conditions (1.12) and (1.14) are operative. Then for any $r > 0$ and $q \in \mathbb{N}$ such that $\gamma^2 q < 2$, we have*

$$\begin{aligned}
 & \lim_{N \rightarrow \infty} L(N)^{-\gamma^2 \binom{q}{2}} \mathbb{E} \left[(\mu_{N, \epsilon_N}^{\gamma, \text{CUE}}([0, r]))^q \right] \\
 &= \lim_{\ell \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \ell^{-\gamma^2 \binom{q}{2}} \mathbb{E} \left[(\mu_{N, \epsilon}^{\gamma, \text{CUE}}[0, r])^q \right] \\
 &= \int_{[0, r]^q} |\Delta(\mathbf{u})|^{-\gamma^2} du_1 \dots du_q \\
 &= r^{q+\gamma^2 \binom{q}{2}} \prod_{j=0}^{q-1} \frac{\Gamma(1 - (j+1)\gamma^2/2) \Gamma^2(1 - j\gamma^2/2)}{\Gamma(1 - \gamma^2/2) \Gamma(2 - (q+j-1)\gamma^2/2)}, \tag{1.15}
 \end{aligned}$$

where $\Delta(\mathbf{u}) = \prod_{1 \leq j < k \leq q} (u_k - u_j)$.

This theorem may be viewed as a smoothed analogue of Conjecture 1.1 and solves the case $q \in \mathbb{N}$ of the conjectures posed by the second author in [55]. In the latter paper, the double limit $N \rightarrow \infty$ followed by $\epsilon \rightarrow 0$ is referred to as the *weak conjecture*, while the more fundamental single limit in the first equality is the *strong conjecture*. Theorem 1.6 demonstrates the equivalence of the strong and weak limits for positive integer $q \in \mathbb{N}$. The full statement of these conjectures consist of assertions valid for any $q \in \mathbb{C}$; such generality is beyond the scope of this paper, though one expects Theorem 1.6 to hold for general $q \in \mathbb{C}$ with the appropriate analytic continuation of the Selberg integral as explained below, *c.f.* equation (49) in [55]. The exponent of the factor $r^{q+\gamma^2 \binom{q}{2}}$ in (1.15) is known as the *structure exponent* and its quadratic dependence on q are associated with the multi-fractal nature of the underlying random measures [60]. It is natural to ask what happens if we take a different interval in (1.6); indeed in [55] the examples $[0, u]$ and $[-L(N), u]$ are studied in detail and the analogous results conjectured there for $q \in \mathbb{N}$ follow easily with our approach. In fact the second example $[-L(N), u]$ behaves similarly to our interval $[u - L(N)/2, u + L(N)/2]$ with $L(N) \rightarrow \infty$. However, it turns out that the random measures corresponding to these examples are *not normalizable* in the usual GMC sense of (1.11). We leave it as an interesting open problem to find a mathematically rigorous way of normalizing the GMC measures corresponding to these interval statistics.

The integral on the second line of (1.15) is a particular case of a multi-dimensional integral known as *Selberg's integral*, due to its explicit evaluation by Selberg in 1944, see [24] for a detailed historical review (similarly the integral in (1.3) is referred to as the *Dyson integral*). It is known that the Dyson and Selberg integrals describe the moments of the total mass of the Bacry-Muzy Gaussian multiplicative chaos measure on the circle and the interval, respectively. Hence the problem of extending these integrals to meromorphic functions of $q \in \mathbb{C}$ in such a way that the resulting function is the Mellin transform of a probability distribution is of fundamental importance as the probability distribution is then naturally conjectured to be that of the total mass. This problem was solved by Fyodorov and Bouchaud [25] for the Dyson integral, heuristically by Fyodorov *et. al.* [30] and independently and rigorously by the second author [52, 53, 56] for the Selberg integral, see also [54] for the Morris integral. These analytic extensions are fundamental in the theory of log-correlated Gaussian processes [25–27, 29–31] as they lead to precise conjectures about the explicit form of the extreme value statistics, by an analogy with the so-called derivative martingale coming from branching processes. The

connection between the derivative martingale and the distribution of the maximum has been rigorously proved for a certain class of log-correlated Gaussian fields [21].

In view of our Theorems 1.3 and 1.6, it is reasonable to ask whether they tell us anything new about eigenvalue statistics of the CUE. It turns out that Theorem 1.3 can be used to obtain a lower bound on the maximum of the process $W_N(u)$, (1.6). Namely, by first proving a lower bound for the maximum of the field $\bar{X}_{N,\epsilon_N}(u)$ where $\epsilon_N = N^{\alpha-1+\kappa}$ and $\kappa > 0$ is small and then by taking care of removing the regularization, it can be shown that

$$\lim_{N \rightarrow \infty} \mathbb{P} \left[\max_{|u| \leq C} W_N(u) \geq \frac{\ell}{2} N^{1-\alpha} + (\sqrt{2} - \delta) \log N^{1-\alpha} \right] = 1, \quad (1.16)$$

for any mesoscopic scale $0 < \alpha < 1$, any small $\delta > 0$, and for any large $C > 0$. Note that $\mathbb{E}[W_N(u)] = \frac{\ell}{2} N^{1-\alpha}$ and the constant $\sqrt{2}$ corresponds to the GMC critical value in dimension 1. In the global regime, our method produces the same bound with $\alpha = 0$ and $C = \pi$, providing a result analogous to [1, Theorem 1.2] about the leading order of the maximum of the imaginary part of the characteristic polynomial. In this case, it is known that the GMC critical value describes the leading order of the maximum of the field in the sense that for any $\delta > 0$,

$$\lim_{N \rightarrow \infty} \mathbb{P} \left[\max_{|u| \leq \pi} W_N(u) \leq \frac{\ell}{2} N^{1-\alpha} + (\sqrt{2} + \delta) \log N \right] = 1. \quad (1.17)$$

In general, it is a difficult task to establish a lower bound for the extreme values of log-correlated fields. On the other hand, we will prove in a future publication that the bound (1.16) follows from the convergence (1.13) provided that it holds throughout the subcritical phase. The main idea is simple and relies on the fact that the random measure $\mu_{N,\epsilon_N}^{\gamma, \text{CUE}}$ lives on the set of γ -thick points:

$$\mathcal{T}_N^\gamma := \left\{ x \in \mathfrak{A} : \bar{X}_{N,\epsilon_N}(x) \geq \gamma \mathbb{E} \left[\bar{X}_{N,\epsilon_N}(x)^2 \right] \right\}, \quad (1.18)$$

for large N , so that if its limit is non-trivial, there must exist (random) points where the field takes atypically large values. It is an interesting question whether the convergence of the exponential measure in the critical case, $\gamma = \sqrt{2d}$, would yield some further information about extreme value statistics and whether it is possible to extend the validity of Theorem 1.3 to this case.

An interesting consequence of the lower-bound (1.16) concerning eigenvalue rigidity⁴ is that for any $\delta > 0$ and scale $\alpha < 1$, when N is large there exists an arc A of size $N^{-\alpha}$ so that

$$\#\{\theta_j \in A\} \geq \mathbb{E}[\#\{\theta_j \in A\}] + (\sqrt{2} - \delta) \log(N|A|), \quad (1.19)$$

with high probability. That is the eigenvalues are over-crowded in this arc, since one typically expects the second term in (1.19) to be of order $\sqrt{\log N}$. We are planning in future work to further discuss the consequences of our results for extreme value statistics of log-correlated fields and eigenvalue rigidity in the random matrix context. In particular, for the sine process using the results of Sect. 1.3.

⁴ By rigidity, we mean that the number of CUE eigenvalues contained in any (mesoscopic) arc A on the unit circle concentrates near its mean $|A| \frac{N}{2\pi}$ with high probability when the dimension N is large.

1.2. Strategy of the proof. In order to prove Theorem 1.3, we develop a general method for constructing multiplicative chaos measures in models where the log-correlated fields are only asymptotically Gaussian. In general, this requires asymptotics for the exponential moments of the field, which is much stronger than the usual approximation given by a CLT such as (1.8). The precise assumptions of our theory are rather technical and we present them in full detail in Sect. 2 together with the proof of Theorem 1.7 below. In this introduction, for the sake of transparency, we formulate our results in a simpler setting which is closer to that described in Sect. 1.1, but remains rather general.

Theorem 1.7. *As in Theorem 1.2, let G be a log-correlated Gaussian field with covariance (1.4), ϕ be a mollifier, and $G_\epsilon = G * \phi_\epsilon$ for any $\epsilon > 0$. For each $N \in \mathbb{N}$, let $(X_{N,\epsilon}(u))_{u \in \mathfrak{A}, \epsilon \in (0,1]}$ be a centered random field on $\mathfrak{A} \subset \mathbb{R}^d$ (not necessarily Gaussian). Suppose that there exists a sequence $\delta_N \rightarrow 0$ as $N \rightarrow \infty$ so that for any $q \in \mathbb{N}$ and $\gamma \in \mathbb{R}^q$,*

$$\log \mathbb{E} \left[\exp \left(\sum_{j=1}^q \gamma_j X_{N,\epsilon_j}(u_j) \right) \right] = \frac{1}{2} \sum_{j,k=1}^q \gamma_j \gamma_k \mathbb{E} [G_{\epsilon_j}(u_j) G_{\epsilon_k}(u_k)] + o(1) \quad (1.20)$$

$N \rightarrow \infty$

uniformly for all $\mathbf{u} \in \mathfrak{A}^q$ and $\boldsymbol{\epsilon} \in (\delta_N, 1]^q$. Then, for any $0 < \gamma < \sqrt{2d}$ and $w \in L^1(\mathfrak{A})$ uniformly bounded, the sequence of random variables

$$\mu_N^\gamma(w) := \int e^{\gamma X_{N,\delta_N}(u) - \frac{\gamma^2}{2} \mathbb{E}[X_{N,\delta_N}(u)^2]} w(u) du \quad (1.21)$$

converges in distribution to the random variable $v^\gamma(w)$ where v^γ is the GMC measure defined in Theorem 1.2 associated with the field G .

Let us note that in the above, the field $X_{N,\epsilon}(u)$ need not be defined on the same probability space for different N . Theorem 1.7 establishes the convergence to a GMC measure in the whole subcritical phase $\gamma < \sqrt{2d}$. The proof is inspired by the elementary argument introduced by Berestycki [5] for establishing Theorem 1.2. Our main contributions are to single out the assumptions needed to apply these ideas in a setting where the fields are only asymptotically Gaussian and find a way to replace the Gaussian techniques (like, for instance, the use of Girsanov's theorem) with other estimates using the strong asymptotics⁵ (1.20). The strategy of the proof of Theorem 1.7 consists in constructing a (random) set $\mathfrak{S} \subset \mathfrak{A}$ which depends on the parameters $0 < \gamma < \sqrt{2d}$ and $N \in \mathbb{N}$ in such a way that $\mu_N^\gamma(w \mathbb{1}_{\mathfrak{S}})$ is uniformly bounded in $L^2(\mathbb{P})$ and $\lim_{N \rightarrow \infty} \mathbb{E} [\mu_N^\gamma(w \mathbb{1}_{\mathfrak{A} \setminus \mathfrak{S}})] = 0$. This second moment computation makes essential use of the uniformity of the error term in the asymptotics (1.20). These properties show that the sequence of random variables $\mu_N^\gamma(w)$ is tight and we identify its limit by showing that

$$\lim_{N \rightarrow \infty} \mu_N^\gamma(w) \stackrel{d}{=} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mu_{N,\epsilon}^\gamma(w) \stackrel{d}{=} v^\gamma(w).$$

By “ $\stackrel{d}{=}$ ” we mean that these random variables have the same law. This last idea is taken from the work of Webb, [70], on the convergence of the characteristic polynomial of the CUE. However, Webb only established the convergence in the L^2 -phase, $\gamma < \sqrt{d}$, by showing that in this regime, $\mu_N^\gamma(w)$ is uniformly bounded in $L^2(\mathbb{P})$ and is

⁵ In fact, the argument presented in Sect. 2 does not require that the asymptotics (1.20) hold in full generality.

therefore uniformly integrable—see Sect. 2.1. In the so-called L^1 -phase, $\sqrt{d} < \gamma < \sqrt{2d}$, $\mathbb{E}[\mu_N^\gamma(1)^2] \rightarrow \infty$ as $N \rightarrow \infty$ and, in effect, we need to restrict the random measure μ_N^γ to the set of points which are *not* α -thick, see (1.18), for some parameter α slightly bigger than γ , to restore the uniform integrability property. This is the main idea to construct the set \mathfrak{S} .

As we already mentioned, Theorem 1.3 follows from an application of Theorem 1.7 to the random field (1.7) where G is a stationary Gaussian field on \mathbb{R} with covariance structure:

$$\mathbb{E}[G(u)G(v)] = Q(u - v) = \log \frac{1}{|u - v|} + \log |\ell^2 - (u - v)^2|^{1/2}, \quad (1.22)$$

see formula (1.10). For the CUE, since the random variables $X_{N,\epsilon}(u)$ are linear statistics, the strong Gaussian asymptotics (1.20) can be checked by exploiting the connection between the CUE and Toeplitz or Fredholm determinants. In particular, we use the following remarkable formula.

Theorem 1.8 (Borodin–Okounkov–Case–Geronimo formula). *Let g be a function on the unit circle satisfying the regularity condition*

$$\sigma^2(g) := \sum_{k=1}^{\infty} k |\hat{g}(k)|^2 < \infty, \quad \text{where} \quad \hat{g}(k) = \frac{1}{2\pi} \int_0^{2\pi} g(e^{i\theta}) e^{-ik\theta} d\theta. \quad (1.23)$$

Then we have the exact formula

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^N g(e^{i\theta_j}) \right) \right] = \exp \left(N \hat{g}(0) + \frac{1}{2} \sigma^2(g) \right) \det (1 + R_N V(g) V(g)^* R_N) \quad (1.24)$$

where $V(g)$ is a certain Hilbert–Schmidt operator acting on $\ell^2(\mathbb{N}, \mathbb{C})$ (see Sect. 3.2 for full details) and R_N is the orthogonal projection onto the subspace $\ell^2(\{0, \dots, N-1\}, \mathbb{C})$.

For a comprehensive account of this formula, which originally appeared in [33], see the monograph of Simon [67]. The first factor on the RHS of formula (1.24) corresponds to the Laplace transform of a Gaussian random variable with mean $N \hat{g}(0)$ and variance $\sigma^2(g)$. The second factor is a Fredholm determinant of an operator acting on the sequence space $\ell^2(\mathbb{N}, \mathbb{C})$. For fixed g , the operator $R_N V(g) V(g)^* R_N$ converges to zero in the trace norm which implies the Strong Szegő limit theorem. In the context of obtaining estimate (1.20), we have⁶

$$\sum_{j=1}^q \gamma_j X_{N,\epsilon_j}^{(2\pi)}(u_j) = \sum_{j=1}^N g_N(e^{i\theta_j})$$

where g_N is a smooth function on the unit circle varying with N . Hence, in order to obtain the asymptotics (1.20), the main challenge is to show that, uniformly in the various parameters, $\lim_{N \rightarrow \infty} \det(1 + R_N V(g_N) V(g_N)^* R_N) = 1$. This is precisely where the condition (1.12) comes into play. Namely, there is a transition in the asymptotics when the regularization scale ϵ_N becomes of the order of the mean eigenvalue spacing $N^{\alpha-1}$.

⁶ Here $X_{N,\epsilon}^{(2\pi)}$ is an appropriate periodization of $X_{N,\epsilon}$ which does not change the Gaussian asymptotics (1.20), see Sect. 3.2.

When $\epsilon_N = O(N^{\alpha-1})$, the Fredholm determinant on the RHS of formula (1.24) does not converge to 1 and we enter the so-called Fisher–Hartwig regime characterized by the case $\epsilon_N \equiv 0$. In fact, if we consider the counting statistics (1.6) and set $\overline{W}_N = W_N - \mathbb{E}W_N$, we deduce from [44, Theorem 2.2] that for any $q \in \mathbb{N}$ and $\gamma \in (-1, \infty]^q$,

$$\begin{aligned} \mathbb{E} \left[\exp \left(\sum_{j=1}^q \gamma_j \overline{W}_N(u_j) \right) \right] &= \exp \left(\frac{1}{2} \sum_{j=1}^q \gamma_j^2 \log N - \sum_{j < k} \gamma_j \gamma_k \log |e^{iu_k} - e^{iu_j}| \right. \\ &\quad \left. + \sum_{j < k} \frac{\gamma_j \gamma_k}{2} \log \left(|e^{i(u_k - \ell)} - e^{iu_j}| |e^{i(u_k + \ell)} - e^{iu_j}| \right) \right) \prod_{j=1}^q \left| G(1 + \frac{\gamma_j}{2i}) \right|^4 (1 + o(1)), \end{aligned} \quad (1.25)$$

for fixed parameters $0 < u_1 < \dots < u_q$ (the error term is no longer uniform). For the proof of Theorem 1.7, it is crucial that the asymptotics (1.20) are uniform when the points u_j merge. When $q = 2$ this merging has been studied by Claeys and Krasovsky [14] and very precise asymptotics are known in the various regimes.

However, for $q > 2$, there are no results about merging singularities and this is known to be a complicated problem. This lack of uniformity is the main obstacle in establishing the analogues of Theorems 1.3 and 1.6 without the regularization procedure ($\epsilon_N = 0$). Therefore, the condition (1.12) simplifies the asymptotics by assuring that the regularized field remains in a Gaussian regime and prevents the technical complications due to the emergence of Fisher–Hartwig singularities. Nevertheless, as (1.16) shows, such a condition still allows us to recover a sharp lower bound for the leading order behavior of the maximum of the field.

Finally, let us remark that the Barnes G-functions in (1.25) seem to indicate non-Gaussianity of the process \overline{W}_N beyond the leading order for large N . However, we expect that for a fixed γ the normalization used by Webb [70] $\frac{e^{\gamma \overline{W}_N(u)}}{\mathbb{E}(e^{\gamma \overline{W}_N(u)})}$ and the usual normalization $e^{\gamma \overline{W}_N(u) - \frac{\gamma^2}{2} \mathbb{E}(\overline{W}_N(u)^2)}$ converge to the same limiting random variable up to a constant factor depending only on γ .

1.3. Results for the sine process. In this subsection, we explain how to apply Theorem 1.7 to regularized counting statistics of the sine process. The results are analogous to those stated for the CUE in Sect. 1.1. Hence, we expect that similar results hold for well-behaved unitary invariant ensembles as well, but we leave the task of proving the sufficient asymptotics open for a future project. The sine process, denoted by Λ_N , is the determinantal point process on \mathbb{R} with correlation kernel

$$K_N(u, v) = \frac{\sin(\pi N(u - v))}{\pi(u - v)}. \quad (1.26)$$

We refer to Sect. 1.5 below for some background on determinantal processes. This is a translation invariant point process whose density is N times the Lebesgue measure on \mathbb{R} . Recall that, given a mollifier ϕ , we denote $\phi_\epsilon(x) = \epsilon^{-1} \phi(x/\epsilon)$ and $\chi_u(x) = \pi \mathbb{1}_{|x-u| \leq \ell/2}$ for all $u, x \in \mathbb{R}$. Let us consider the linear statistics:

$$X_{N, \epsilon}(u) = \sum_{\lambda \in \Lambda_1} \chi_u * \phi_\epsilon(\lambda/N). \quad (1.27)$$

As $N \rightarrow \infty$, the random variable (1.27) gives an approximation of the number of eigenvalues in a mesoscopic box in the bulk of the GUE, or say of another unitary invariant ensemble. To see the parallel with the CUE, note that the scaling property of the sine kernel implies that

$$X_{N,\epsilon}(u) \stackrel{d}{=} \sum_{\lambda \in \Lambda_N} \chi_u * \phi_\epsilon(\lambda). \quad (1.28)$$

We will focus on the strong regime where $\epsilon = \epsilon_N \rightarrow 0$ as $N \rightarrow \infty$ and, viewing $X_{N,\epsilon}$ as an asymptotically Gaussian field on \mathbb{R} , we will construct its chaos measure in the subcritical phase. The advantage of working with the random variables (1.27) instead of the RHS of (1.28) is that it introduces a natural coupling which allows us to obtain a stronger mode of convergence than for the CUE. For technical reasons, we will restrict ourselves to the following class of real-analytic mollifiers.

Assumption 1.1. Suppose that the function ϕ is analytic in $|\Im z| < c$ and that $|\Im \phi| < \pi/\ell$ in this strip. We also assume that $\phi \geq 0$ on \mathbb{R} , $\int_{\mathbb{R}} \phi(x) dx = 1$, and $\sup \{ \|\phi\|_{L^1(\mathbb{R}+is)} : s \leq c/2 \} = C_\phi < \infty$.

For any $\gamma > 0$, we now consider the random measure:

$$\mu_{N,\epsilon}^{\gamma, \text{Sine}}(w) = \int_{\mathbb{R}} e^{\gamma \bar{X}_{N,\epsilon}(u) - \frac{\gamma^2}{2} \mathbb{E}(\bar{X}_{N,\epsilon})^2} w(u) du, \quad (1.29)$$

where, as for the CUE, $\bar{X}_{N,\epsilon} := X_{N,\epsilon} - \mathbb{E}X_{N,\epsilon}$. We also define a space of mollifiers that will enter into the regularized field (1.28): For any $\alpha \geq 0$, let

$$\mathcal{D}_\alpha = \left\{ \phi \in L^1 \cap L^2(\mathbb{R}) : \phi \geq 0, \int_{\mathbb{R}} \phi(x) dx = 1 \text{ and } \int_{\mathbb{R}} |x|^\alpha \phi(x) dx < \infty \right\} \quad (1.30)$$

and $\mathcal{D} = \bigcup_{\alpha > 0} \mathcal{D}_\alpha$. For instance, the functions $\phi(z) = e^{-z^2/2}/\sqrt{2\pi}$ or $\phi(z) = \frac{1/\pi}{1+z^2}$ satisfy Assumption 1.1 and belong to the set \mathcal{D} . Finally recall that G is the stationary Gaussian field on \mathbb{R} with zero mean and covariance kernel (1.22). As in the case of the CUE, we let ν^γ be the GMC measure associated to the field G .

Theorem 1.9. Let $w \in L^1 \cap L^\infty(\mathbb{R})$, $\phi \in \mathcal{D}$ be a function which satisfies Assumption 1.1, and let ϵ_N be a sequence which converges to 0 as $N \rightarrow \infty$ in such a way that $\epsilon_N \geq N^{-1}(\log N)^{1+\kappa}$ for some $\kappa > 0$. For any $0 < \gamma < \sqrt{2}$, $\mu_{N,\epsilon_N}^{\gamma, \text{Sine}}(w)$ converges in $L^1(\mathbb{P})$ as $N \rightarrow \infty$ to a random variable $\mu^\gamma(w)$ which has the same law as $\nu^\gamma(w)$, where ν^γ is the GMC measure defined in Theorem 1.2 and $g(u, v) = \log |\ell^2 - (u - v)^2|^{1/2}$.

This shows that, in the subcritical phase, the law of the random measure μ^γ does not depend on the mollifier ϕ and it is the same GMC measure as for the CUE and as for Theorem 1.2. The proof of Theorem 1.9 is given at the end of Sect. 2.2 and is a direct consequence of our general result, Theorem 2.6. The main assumption to obtain Theorem 1.9 is again the strong Gaussian approximation (1.20). The needed exponential moments of (1.28) can be expressed in terms of a Fredholm determinant; see formula (1.39) below. In Sect. 1.5, we explain how the asymptotics of this Fredholm determinant can be related to a 2×2 Riemann–Hilbert problem. Then, the condition $\epsilon_N \geq N^{-1}(\log N)^{1+\kappa}$ guarantees that we can straightforwardly apply the Deift–Zhou steepest descent method to this problem.

Assumption 1.2. Let $0 < \alpha < \pi$. Suppose that h is a function which is analytic and satisfies

$$|\Im h(z)| < \alpha,$$

in a strip $|\Im z| < \delta$. We also assume that $h : \mathbb{R} \rightarrow \mathbb{R}$, that $h' \in L^1(\mathbb{R})$, and that the following constants are finite:

$$\begin{aligned} C_\infty &= \sup \left\{ \exp |h(z)| : |\Im z| \leq \delta/2 \right\} \text{ and } C_1 \\ &= \sup \left\{ \|h\|_{L^\infty(\mathbb{R}+is)} \vee \|h\|_{L^1(\mathbb{R}+is)} : s \leq \delta/2 \right\}. \end{aligned} \quad (1.31)$$

We need to work with real-analytic mollifiers in order to apply the Riemann–Hilbert machinery. In principle, it could be possible to work with a more general class of mollifiers by using an argument analogous to the one in [7]. It is based on constructing an N -dependent approximation $\phi^{(N)}$ of the mollifier ϕ which is real-analytic so that we can solve the Riemann–Hilbert problem for $\phi^{(N)}$, c.f. Lemma 3.1, and argue that the Laplace transform of the two regularizations are sufficiently close as $N \rightarrow \infty$.

Proposition 1.10. *If h is a test function which satisfies Assumption 1.2, then*

$$\log \mathbb{E} \left[\exp \left(\sum_{\lambda \in \Lambda_N} h(\lambda) \right) \right] = N \int h(x) dx + \frac{1}{2} \|h\|_{H^{1/2}(\mathbb{R})}^2 + O_{N \rightarrow \infty} \left(\frac{C_\infty^6 C_1^2}{\delta^{3/2} |\sin \alpha|} e^{-\pi \delta N} \right). \quad (1.32)$$

where $\|h\|_{H^{1/2}(\mathbb{R})}^2$ is defined in Appendix A, equation (A.8). The implied constant in the big- O term of (1.32) does not depend on h , δ or N .

The proof of Proposition 1.10 is given in Sect. 3.1. This is a classical result, [17], but we need to take extra care to control the error term uniformly, especially because we will consider N -dependent test functions. Specifically, we can consider any regime where $\delta(N) \rightarrow 0$ as $N \rightarrow \infty$ almost as fast as N^{-1} (i.e. up to the critical regime where the Gaussian approximation fails). In fact, our asymptotics are sufficiently strong to strengthen the convergence of the measure $\mu_{N,\epsilon}^{\gamma, \text{Sine}}$ when the parameter γ is sufficiently small. In particular, motivated by the conjectures of [55], beyond the L^2 -phase, we establish the convergence of all the existing moments of the multiplicative chaos measure $\mu_{N,\epsilon}^{\gamma, \text{Sine}}$.

Theorem 1.11. *Under the same assumptions as Theorem 1.9, if q is an even integer and $\gamma \geq 0$ so that $q\gamma^2 < 2$, then the random variable $\mu_{N,\epsilon}^\gamma(w)$ converges in $L^q(\mathbb{P})$ to $\mu^\gamma(w)$. Moreover, for any $q \in \mathbb{N}$ and $\gamma \geq 0$ such that $q\gamma^2 < 2$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\mu_{N,\epsilon}^{\gamma, \text{Sine}}(w)^q \right] = \int_{\mathbb{R}^q} \exp \left(\gamma^2 \sum_{1 \leq j < k \leq q} Q(u_j - u_k) \right) \prod_{k=1}^q w(u_k) du_k. \quad (1.33)$$

where the correlation kernel Q is given by formula (1.22).

The proofs of Theorems 1.9 and 1.11 are given at the end of Sects. 2.2 and 2.1 respectively. Finally, let us mention that the main challenge to extend Theorem 1.3 beyond the CUE or sine process boils down to obtaining the strong Gaussian asymptotics (1.20). This is an interesting problem, also of independent interest, for both Hermitian unitary invariant or Wigner ensembles.

1.4. Overview of the paper. The paper is organized as follows. In Sect. 1.5, we begin with a brief review of determinantal point processes associated with integrable operators, of which the CUE and sine process are special cases. In Sects. 2.1 and 2.2, we develop the multiplicative chaos theory for random fields which satisfy the strong Gaussian approximation (1.20) and prove Theorem 1.7. In Sect. 2.3, we analyze the covariance structure of the regularized field $G_\epsilon(u)$ arising from the $H^{1/2}$ noise (1.9), proving some preparatory results to apply the general theory to the CUE and sine process. In Sects. 3.1 and 3.2, we establish the required asymptotics (1.20) for the sine and CUE point processes, using a Riemann–Hilbert problem and formula (1.24), respectively. Finally, in the appendix (Sect. B), we provide a review of some of the recent developments in the theory of Gaussian multiplicative chaos of relevance to the present article.

In what follows, $C > 0$ denotes a numerical constant which may change from line to line and we use the notation $a \ll b$ to specify that the quantity $a \leq Cb$. We also define for all $x \in \mathbb{R}$,

$$\log^+(x) = \log(1 \vee |x|), \quad (1.34)$$

where throughout the article we use the notation $x \wedge y := \min\{x, y\}$ and $x \vee y := \max\{x, y\}$.

1.5. Determinantal point processes and integrable operators. The aim of this section is to provide, in a general context, a short introduction to the theory of determinantal point processes which focuses on the connection between linear statistics and Fredholm determinants. We also briefly review the concept of integrable operators introduced in [38] and how this relates the Laplace transform of a linear statistic to a Riemann–Hilbert problem.

Let Σ be a Polish space equipped with a Radon measure η . A point configuration $\Upsilon \subset \Sigma$ is a discrete set which is locally finite (i.e. the set $\Upsilon \cap B$ is finite for any compact set $B \subset \Sigma$). A point process is a probability measure on the space of point configurations. This definition can be made mathematically precise, see for instance [8, 40, 69], and a point process can be described by its intensity measures or correlation functions $\{\rho_n\}_{n=1}^\infty$ which are defined by the formulae:

$$\mathbb{E} \left[\sum_{(\lambda_1, \dots, \lambda_n) \subset \Upsilon} \prod_{k=1}^n f_k(\lambda_k) \right] = \int_{\Sigma^n} \prod_{k=1}^n f_k(x_k) \rho_n(dx_1, \dots, dx_n), \quad (1.35)$$

for any functions $f_1, \dots, f_n \in L^\infty(\Sigma \rightarrow \mathbb{R}_+)$ with compact support. Note that the LHS of formula (1.35) consists of a sum over all ordered subsets of the random configuration Υ of size $n \in \mathbb{N}$. A point process is called determinantal if all its intensity measures are of the form

$$\rho_n(dx_1, \dots, dx_n) = \det_{n \times n} [K(x_i, x_j)] \eta(dx_1) \cdots \eta(dx_n). \quad (1.36)$$

The function $K : \Sigma \times \Sigma \rightarrow \mathbb{C}$ is called the correlation kernel. It is obviously not unique, but it encodes the law of the random configuration Υ . There are many interesting examples of determinantal processes coming from probability theory, combinatorics, and mathematical physics such as the eigenvalues of unitary invariant random matrices, free fermions, zeros of Gaussian analytic functions, non-intersecting random walks, uniform spanning trees, random tilings, etc. We refer to the surveys [8, 35, 40] for further examples. Let us just mention the following criterion which goes back to the beginning of the theory, [48], and describes a natural class of correlation kernels.

Theorem 1.12 (Macchi [48], Soshnikov [69]). *If a kernel K determines a self-adjoint integral operator acting on $L^2(\Sigma)$ which is locally trace-class, then K defines a determinantal point process if and only if its spectrum is contained in $[0, 1]$.*

In the following, we shall assume that the kernel K is a continuous function on $\Sigma \times \Sigma$ and satisfies the hypothesis of Theorem 1.12. In this case, this kernel defines an operator, also denoted by K , which is locally trace-class if and only if for any compact set $B \subseteq \Sigma$,

$$\mathrm{Tr} K = \int_B K(x, x) dx < \infty;$$

see [67, Theorem 2.12]. We let Υ be the point configuration of the determinantal process with kernel K and for any function $\varphi \in L^\infty(\Sigma \rightarrow \mathbb{R}_+)$, we denote

$$K_\varphi(x, x') = \sqrt{\varphi(x)} K(x, x') \sqrt{\varphi(x')}. \quad (1.37)$$

The condition that the function $\varphi \geq 0$ is not necessary but rather convenient. In particular, this implies that the operator K_φ is also self-adjoint, non-negative, and it is trace class if

$$\mathrm{Tr} K_\varphi = \int_\Sigma K(x, x) \varphi(x) dx = \mathbb{E} \left[\sum_{\lambda \in \Upsilon} \varphi(\lambda) \right] < \infty. \quad (1.38)$$

Note that this condition holds if for instance φ has compact support. The last equality in (1.38) follows from the definition of the first intensity measure and the function $x \mapsto K(x, x)$ is called the density of the point process Υ . The reason to consider the kernel (1.37) is that, using formulae (1.35) and (1.36), it is a simple combinatorial exercise to show that if K_φ is trace-class, then for any $t \geq 0$,

$$\mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + t\varphi(\lambda)) \right] = \det[\mathrm{I} + tK_\varphi]_{L^2(\Sigma)}, \quad (1.39)$$

where the RHS is a Fredholm determinant; c.f. [40]. In particular, taking the usual logarithm, we obtain

$$\log \mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + t\varphi(\lambda)) \right] = \mathrm{Tr} \log(\mathrm{I} + tK_\varphi), \quad (1.40)$$

and this function is differentiable for all $t > 0$:

$$\frac{d}{dt} \mathrm{Tr} \log(\mathrm{I} + tK_\varphi) = \mathrm{Tr} \left[\frac{K_\varphi}{\mathrm{I} + tK_\varphi} \right].$$

Hence, if we define $L_t := \frac{K_\varphi}{\mathrm{I} + tK_\varphi}$, this implies that

$$\log \mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + \varphi(\lambda)) \right] = \int_0^1 \mathrm{Tr}[L_t] dt. \quad (1.41)$$

For instance, taking $\varphi(x) = \mathbb{1}_{x \in B}$ for some compact subset $B \subseteq \Sigma$, one can investigate the distribution of the random variable $|\Upsilon \cap B|$ and in particular the probability that there are no points in the set B . More generally, if $h \in L^\infty(\Sigma \rightarrow \mathbb{R}_+)$, taking

$\varphi(x) = e^{h(x)} - 1$, this gives an explicit formula for the exponential moments or Laplace transform of the linear statistic $\sum_{\lambda \in \Lambda} h(\lambda)$. As the density of the point process converges to infinity, this reduces the question about the statistical properties of the random variable $\sum_{\lambda \in \Upsilon} h(\lambda)$ to a question about the asymptotics of the resolvent operator L_t . There is a special class of determinantal processes, those for which the correlation kernel gives rise to an integrable operator, which are particularly interesting because computing the resolvent L_t turns out to be equivalent to solving a Riemann–Hilbert problem; see Proposition 1.13 below. In particular, it allows to use the so-called *Deift–Zhou steepest descent* method introduced in [20] to obtain the asymptotics of formula (1.41). The theory of integrable operators and the auxiliary Riemann–Hilbert problem originates in the context of statistical field theory [38], but this approach has also been used to answer different types of questions about the statistics of eigenvalues of unitary invariant matrix ensembles. For instance, one can find a proof of the Strong Szegő limit theorem in [17] and, in [7], the authors extended Deift’s method to investigate a transition for smooth mesoscopic statistics of the so-called thinned CUE and thinned sine process. Mesoscopic statistics were also studied using a Riemann–Hilbert problem in [28].

In this paper, we will use an analogous method to derive the necessary estimates to construct a multiplicative chaos measure which arises naturally from the sine process. In particular, we will make use of the following result from [38], see also [17].

Theorem 1.13. *Suppose that Σ is a closed (oriented) curve on the Riemann sphere. Let Υ be a determinantal process on Σ with Hermitian correlation kernel of the form*

$$K(z, z') = -\frac{f(z)^* g(z')}{2\pi i (z - z')}, \quad (1.42)$$

where $f : \Sigma \rightarrow \mathbb{C}^k$ and $g : \Sigma \rightarrow \mathbb{C}^k$ are continuously differentiable functions so that $f(z)^* g(z) = 0$ for all $z \in \Sigma$. If $\varphi : L^\infty(\Sigma \rightarrow \mathbb{R})$ is a test function so that both $\sqrt{\varphi}f, \sqrt{\varphi}g \in L^2(\Sigma)$, then we have

$$\log \mathbb{E} \left[\prod_{\lambda \in \Upsilon} (1 + \varphi(\lambda)) \right] = \frac{-1}{2\pi i} \int_0^1 \int_{\mathbb{R}} \left(\frac{d\sqrt{\varphi}F_t}{dx}(x) \right)^* \left(\sqrt{\varphi(x)} G_t(x) \right) dx dt \quad (1.43)$$

where $F_t = m_+ f$ and $G_t = (m_+^{-1})^* g$ and the matrix m is the (unique) solution of the Riemann–Hilbert problem:

- $m(z)$ is analytic on $\mathbb{C} \setminus \Sigma$.
- If we let $v = I + t\varphi f g^*$, then $m(z)$ satisfies the jump condition:

$$m_+(z) = m_-(z)v(z), \quad z \in \Sigma. \quad (1.44)$$

- $m(z) \rightarrow I$ as $z \rightarrow \infty$ in $\mathbb{C} \setminus \Sigma$.

For instance, when $\Sigma = \mathbb{R}$, $m_\pm(x) = \lim_{\delta \rightarrow 0} m(x + i\delta)$ denotes the boundary value of the matrix m and are typically assumed to be continuous functions. In practice, $\Sigma = \{|z| = 1\}$ for the CUE, or $\Sigma = \mathbb{R}$ for the sine process and the eigenvalue processes coming from unitary invariant ensembles of Hermitian matrices. Moreover, the correlation kernels of these processes all give rise to integrable operators. According to formula (1.42), we may choose for the CUE:

$$f(z) = \begin{pmatrix} z^{N+1} \\ 1 \end{pmatrix} \quad \text{and} \quad g(z) = \begin{pmatrix} z^{N+1} \\ -1 \end{pmatrix}, \quad |z| = 1$$

and for the sine process:

$$f(x) = \begin{pmatrix} e^{i\pi Nx} \\ e^{-i\pi Nx} \end{pmatrix} \quad \text{and} \quad g(x) = \begin{pmatrix} e^{i\pi Nx} \\ -e^{-i\pi Nx} \end{pmatrix}, \quad x \in \mathbb{R}.$$

For this example, the Riemann–Hilbert problem (1.44) is solved in Sect. 3.1 for a large class of analytic test functions φ ; see in particular Proposition 3.1 for the asymptotics of the solution. As a last comment about universality, if one considers the eigenvalues of an $N \times N$ Hermitian random matrix sampled according to the weight $e^{-N \text{Tr } V(H)}$ for a real-analytic external field $V : \mathbb{R} \rightarrow \mathbb{R}$, then, by the Christoffel–Darboux formula, the correlation kernel of the eigenvalue process is also integrable with

$$\begin{aligned} f_V(x) &= \begin{pmatrix} \pi_N(x) \\ -2\pi i \gamma_{N-1}^2 \pi_{N-1}(x) \end{pmatrix} e^{-NV(x)/2} \quad \text{and} \\ g_V(y) &= \begin{pmatrix} -2\pi i \gamma_{N-1}^2 \pi_{N-1}(x) \\ \pi_N(x) \end{pmatrix} e^{-NV(x)/2}. \end{aligned} \quad (1.45)$$

where π_N and π_{N-1} are the monic polynomials with respect to the measure $e^{-NV(x)} dx$ on \mathbb{R} of degree N and $N - 1$ respectively and

$$\gamma_{N-1}^{-2} = \int_{\mathbb{R}} \pi_{N-1}(x)^2 e^{-NV(x)} dx.$$

Observe that, apart from the weight $e^{-NV(x)/2}$, f_V is exactly the first column of the solution Y_N of the orthogonal polynomial Riemann–Hilbert problem, whose solution is derived in great detail in [19]. In particular from their results, one can extract the universal oscillatory behavior of the functions f_V and g_V in the bulk, which indicates strongly that the approach we present for the sine process could be generalized and would provide a way to show that the limiting chaos measure has the same law for a large class of potentials V . However, turning this heuristic into a rigorous computation is rather technical and we leave it as an open problem for future work.

2. Proof of the Main Results

Even though the applications discussed in this paper are concerned with random processes defined on \mathbb{R} , we will formulate our convergence results in an abstract setting under some general assumptions. Let \mathfrak{A} be a compact set in \mathbb{R}^d , $d \geq 1$. In certain cases, such as the sine process, one might also consider the case where \mathfrak{A} is not compact, this requires only slight modifications of our proof. We consider a real-valued generalized Gaussian process G defined on \mathfrak{A} with a covariance kernel:

$$T(x, y) = -\log |x - y| + g(x, y),$$

where the function $g : \mathfrak{A}^2 \rightarrow \mathbb{R} \cup \{-\infty\}$ is continuous and such that there exists a constant $C > 0$ so that for all $x, y \in \mathfrak{A}$,

$$g(x, y) - \log^+(x - y) \leq C.$$

We also consider a family of real-valued random fields $X_{N,\epsilon}(u)$ defined on \mathfrak{A} which are centered, depend on two parameters $N, \epsilon > 0$, and behave asymptotically like the Gaussian process G . Specifically, we should assume that for any $N > 0$, $(u, \epsilon \mapsto$

$X_{N,\epsilon}(u)_{u \in \mathfrak{A}, \epsilon > 0}$ are random processes defined on the same probability space and that they satisfy Assumptions 2.1–2.4 below. For any $\gamma \in \mathbb{R}$, we consider the normalized process

$$\tilde{X}_{N,\epsilon}^\gamma(u) = \gamma X_{N,\epsilon}(u) - \frac{\gamma^2}{2} \mathbb{E} \left[X_{N,\epsilon}(u)^2 \right], \quad (2.1)$$

and our goal is to construct the limit of the random measure

$$\mu_{N,\epsilon}^\gamma(du) = \exp(\tilde{X}_{N,\epsilon}^\gamma(u)) du$$

when $\epsilon_N \rightarrow 0$ as $N \rightarrow \infty$ sufficiently slowly. To begin with, in Sect. 2.1, we shall prove that $\mu_{N,\epsilon}^\gamma$ converges to a GMC measure in the L^2 -phase ($\gamma < \sqrt{d}$) and compute the limit of its moments in view of proving Theorem 1.6. Then, in Sect. 2.2, we tackle the more challenging task of showing that $\mu_{N,\epsilon}^\gamma$ converges in the whole subcritical regime ($\gamma < \sqrt{2d}$), thus establishing the proof of Theorem 1.7.

2.1. Convergence of the multiplicative chaos measure in the L^2 -phase.

Assumption 2.1 (*Finite-dimensional convergence in the weak regime*). For any given $\epsilon, \delta > 0$, we have

$$\lim_{N \rightarrow \infty} \mathbb{E} [X_{N,\epsilon}(u) X_{N,\delta}(v)] = T_{\epsilon,\delta}(u, v),$$

and the field $u \mapsto X_{N,\epsilon}(u)$ converges in the sense of finite-dimensional distributions to a mean-zero Gaussian process G_ϵ with covariance structure:

$$T_{\epsilon,\delta}(u, v) = \mathbb{E} [G_\epsilon(u) G_\delta(v)], \quad (2.2)$$

for any $u, v \in \mathfrak{A}$ and $\epsilon, \delta > 0$.

In the context of random matrix theory described in the introduction, Assumption 2.1 follows from the CLT for smooth linear statistics and $G_\epsilon = G * \phi_\epsilon$ for some nice mollifier ϕ . In this abstract context, G_ϵ is a $d + 1$ dimensional Gaussian field which is a smooth approximation of a log-correlated G coming from a possibly different regularization procedure. To construct a multiplicative chaos measure out of the field $X_{N,\epsilon}$, one also needs the existence of the GMC measure ν^γ associated with the field G . As discussed in the proof of Proposition B.1 and Remark B.2 below it, the convergence follows from the following conditions for the correlation kernels.

Assumption 2.2 (*Covariance kernel asymptotics*). Suppose that for all $(u, v) \in \mathfrak{A}^2$,

$$T_{\epsilon,\delta}(u, v) \leq \log^+ \left(|u - v|^{-1} \wedge \epsilon^{-1} \wedge \delta^{-1} \right) + C, \quad (2.3)$$

and that for almost all $(u, v) \in \mathfrak{A}^2$,

$$T_{\epsilon,\delta}(u, v) \rightarrow T(u, v) \quad \text{as } \epsilon, \delta \rightarrow 0. \quad (2.4)$$

We also suppose the bound (2.3) is sharp, in the sense that if $\epsilon \geq \delta \geq 0$ and $|u - v| \leq \exp(-\epsilon^{-1})$, then

$$T_{\epsilon,\delta}(u, v) = \log \epsilon^{-1} + O_{\epsilon \rightarrow 0}(1). \quad (2.5)$$

For any $\gamma, \epsilon > 0$, let

$$v_\epsilon^\gamma(dx) = \exp\left(\gamma G_\epsilon(x) - \frac{\gamma^2}{2} \mathbb{E}\left[G_\epsilon(x)^2\right]\right) dx.$$

The Assumption 2.2 guarantees that for any $q \in \mathbb{N}$ such that $\gamma^2 q < 2d$ and for any $w \in L^1 \cap L^\infty(\mathfrak{A})$, the random variable $v_\epsilon^\gamma(w)$ converges in $L^q(\mathbb{P})$ as $\epsilon \rightarrow 0$. The purpose of the next assumption is to identify the limit.

Assumption 2.3 (*Convergence of the GMC measure*). For any $\gamma < \sqrt{2d}$, let v^γ be the GMC measure associated with the field G as defined in Theorem 1.2. Then, $v_\epsilon^\gamma \Rightarrow v^\gamma$ as $\epsilon \rightarrow 0$.

Finally, in order to apply the *second moment method* considered by Webb in [70], we will also need to control some exponential moments of the field $(u, \epsilon) \mapsto \tilde{X}_{N,\epsilon}(u)$. The idea of [70] consists in proving that both in the weak regime (when we consider successive limits as $N \rightarrow \infty$ and $\epsilon \rightarrow 0$) and in the strong regime (when $\epsilon_N \rightarrow 0$ as $N \rightarrow \infty$), the limiting random measures coincide and have the same law as the GMC measure v^γ . In particular, we will need the following asymptotics. For any $q \in \mathbb{N}$ and $\delta > 0$, define

$$\Delta_q(\delta) = \{\epsilon \in \mathbb{R}^q : \epsilon_1 \geq \dots \geq \epsilon_q \geq \delta\}. \quad (2.6)$$

Assumption 2.4 (*Exponential moments asymptotics*). Let $q \in \mathbb{N}$. We suppose that there exists a sequence $\delta_N \rightarrow 0$ as $N \rightarrow \infty$ so that for any $\epsilon \in \Delta_q(\delta_N)$ and for any $\mathbf{t} \in \mathbb{R}^q$, we have uniformly for all $\mathbf{u} \in \mathfrak{A}^q$,

$$\log \mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N,\epsilon_k}^{t_k}(u_k) \right) \right] = \sum_{1 \leq k < j \leq q} t_k t_j T_{\epsilon_j, \epsilon_k}(u_j, u_k) + o(1) \quad \text{as } N \rightarrow \infty. \quad (2.7)$$

In the CUE case (Theorem 1.3), at any mesoscopic scale $0 < \alpha < 1$, one may choose the parameter $\delta_N = N^{\alpha-1+\kappa}$ for some small $\kappa > 0$ so that the condition (1.12) is satisfied. On the other hand, for the sine process with density N , we may choose $\delta_N = N^{-1}(\log N)^{1+\kappa}$ for any $\kappa > 0$.

Theorem 2.1. *Suppose that Assumptions 2.1–2.3 hold, as well as Assumption 2.4 for $q = 1, 2$. If $\gamma^2 < d$, then for any $w \in L^1 \cap L^\infty(\mathfrak{A})$, the random variable $\mu_{N,\delta_N}^\gamma(w)$ converges in distribution as $N \rightarrow \infty$ to $v^\gamma(w)$. Moreover, if Assumption 2.4 is also satisfied for $q \in \mathbb{N}$ and $\gamma^2 q < 2d$, then*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\mu_{N,\delta_N}^\gamma(w)^q \right] = \int \exp \left(\gamma^2 \sum_{1 \leq j < k \leq q} T(u_j, u_k) \right) \prod_{k=1}^q w(u_k) du_k. \quad (2.8)$$

As mentioned in the introduction, the condition $\gamma^2 q < 2d$ for the existence of the limiting moments (2.8) is sharp and is related to the convergence of certain multiple integrals, which in case $d = 1$ are related to Selberg integrals. The remainder of this section is devoted to the proof of Theorem 2.1 and of an extension, Lemma 2.5, in the case where there is a coupling of the fields $X_{N,\epsilon}(u)$ for different $N > 0$. This applies to the sine process discussed in Sect. 1.3. Note that to prove the convergence of the measure μ_{N,δ_N}^γ and its moments in the L^2 -phase, it is clear from our assumptions that one can use the same argument as in the proof of Proposition B.1. However, to identify

that this measure has the same distribution as ν^γ , it is simpler to first establish that $\mu_{N,\epsilon}^\gamma$ converges in distribution in the weak regime and then to show that the strong and weak limits coincide; c.f. Lemmas 2.2 and 2.4 respectively.

Lemma 2.2. *Suppose that \mathfrak{A} is compact in \mathbb{R}^d . Let $w \in L^1(\mathfrak{A})$ and fix $\epsilon > 0$. For any $\gamma \geq 0$, the random variable $\mu_{N,\epsilon}^\gamma(w)$ converges in distribution as $N \rightarrow \infty$ to $\nu_\epsilon^\gamma(w)$.*

Proof. By Assumption 2.1 and continuity of the exponential function, the finite dimensional distributions of the process $\xi_{N,\epsilon}(u) = \exp(\tilde{X}_{N,\epsilon}^\gamma(u))$ converge to those of $\xi_\epsilon(u) = \exp(\tilde{G}_\epsilon^\gamma(u))$ as $N \rightarrow \infty$ where

$$\tilde{G}_\epsilon^\gamma(u) = \gamma G_\epsilon(u) - \frac{\gamma^2}{2} \mathbb{E}(G_\epsilon(u)^2).$$

We also claim that $\xi_{N,\epsilon}(u)$ is tight in $L^1(\mathfrak{A}, |w(u)|du)$, so that, by Prokhorov's theorem, $\xi_{N,\epsilon} \Rightarrow \xi_\epsilon$ as $N \rightarrow \infty$. The tightness follows from a criteria established in [15] which shows that when \mathfrak{A} is compact it suffices that there exists a constant $C_\epsilon > 0$ so that

$$\sup_{u \in \mathfrak{A}} \mathbb{E}[|\xi_{N,\epsilon}(u)|] \leq C_\epsilon. \quad (2.9)$$

Notice that, since G_ϵ is a Gaussian process, $\mathbb{E}[|\xi_\epsilon(u)|] = 1$ and the estimate (2.9) follows directly from Assumption 2.4. Since the functional $\xi \rightarrow \int \xi(u)w(u)du$ is obviously continuous on $L^1(\mathfrak{A}, |w(u)|du)$, we conclude that as $N \rightarrow \infty$,

$$\mu_{N,\epsilon}^\gamma(w) = \int \xi_{N,\epsilon}(u)w(u)du \Rightarrow \nu_\epsilon^\gamma(w) = \int \xi_\epsilon(u)w(u)du.$$

□

Remark 2.3. This proof relies on the fact that the sequence $(\xi_{N,\epsilon})_{N>0}$ is tight in $L^1(\mathfrak{A}, |w(u)|du)$ for any $\epsilon > 0$ and that the condition (2.9) is straightforward to check. However, for the CUE or sine statistics (c.f. (1.7) and (1.27) respectively), using the specific form of the test function $\chi_u * \phi_\epsilon$, it is also possible to verify that the criterion (4) of [43, Theorem 16.5] holds, which implies that $X_{N,\epsilon} \Rightarrow X_\epsilon$ as random elements of $C(\mathfrak{A} \rightarrow \mathbb{R})$.

Lemma 2.4. *Let q be an even integer such that $\gamma^2 q < 2d$. Then, for any $w \in L^1 \cap L^\infty(\mathfrak{A})$,*

$$\lim_{\epsilon \rightarrow 0, N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\delta_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right|^q \right] = 0.$$

Proof. Suppose that $\delta_N \leq \epsilon$. For any $0 \leq i \leq q$, define

$$\epsilon^i = (\underbrace{\epsilon, \dots, \epsilon}_{i\#}, \underbrace{\delta_N, \dots, \delta_N}_{(q-i)\#})$$

and let for any $\mathbf{u} \in \mathfrak{A}^q$,

$$\Theta_{\epsilon,N}^i(\mathbf{u}) = \sum_{1 \leq k < j \leq q} T_{\epsilon_j^i, \epsilon_k^i}(u_j, u_k).$$

By Fubini's theorem and Assumption 2.4 with $\mathbf{t} = (\gamma, \dots, \gamma)$, we obtain

$$\begin{aligned} & \mathbb{E} \left[\left| \mu_{N, \delta_N}^\gamma(w) - \mu_{N, \epsilon}^\gamma(w) \right|^q \right] \\ &= \sum_{i=0}^q (-1)^i \binom{q}{i} \int_{\mathbb{A}^q} \exp \left(\gamma^2 \Theta_{\epsilon, N}^i(\mathbf{u}) + o(1) \right) \prod_{k=1}^q w(u_k) du_k, \end{aligned} \quad (2.10)$$

where the error term is uniform. Moreover, the condition (2.4) implies that for any $i \in [q]$ and for almost all $\mathbf{u} \in \mathbb{A}^q$,

$$\lim_{\epsilon \rightarrow 0, N \rightarrow \infty} \Theta_{\epsilon, N}^i(\mathbf{u}) = \sum_{1 \leq k < j \leq n} T(u_j, u_k).$$

Finally, the condition (2.3) shows that for all $\mathbf{u} \in \mathbb{A}^q$,

$$\exp \left(\gamma^2 \Theta_{\epsilon, N}^i(\mathbf{u}) \right) \ll \prod_{1 \leq j < k \leq q} 1 \vee |u_j - u_k|^{-\gamma^2}. \quad (2.11)$$

Hence, since the RHS of (2.11) is locally integrable on $(\mathbb{R}^d)^q$ when $\gamma^2 q < 2d$, by the dominated convergence theorem, the integrals on the RHS of formula (2.10) converge for all $i \in [q]$ to the same finite value while taking the limit as $N \rightarrow \infty$ and then as $\epsilon \rightarrow 0$. Since $\sum_{i=0}^q (-1)^i \binom{q}{i} = 0$, this proves the claim. \square

Proof of Theorem 2.1. By Lemma 2.2, for any $\epsilon > 0$, we have $\mu_{N, \epsilon}^\gamma(w) \Rightarrow v_\epsilon^\gamma(w)$ as $N \rightarrow \infty$ and Assumption 2.3 guarantees that $v_\epsilon^\gamma(w) \Rightarrow v^\gamma(w)$ as $\epsilon \rightarrow 0$. Hence, by [43, Theorem 4.28] and Lemma 2.4 with $q = 2$, this implies that $\mu_{N, \delta_N}^\gamma(w) \Rightarrow v^\gamma(w)$ as $N \rightarrow \infty$. To complete the proof, it remains to establish convergence of the moments of the random variable $\mu_{N, \delta_N}^\gamma(w)$. We proceed as in the proof of Lemma 2.4. By Fubini's theorem, for any $q \in \mathbb{N}$ and $\epsilon > 0$, we have

$$\mathbb{E} \left[\mu_{N, \epsilon}^\gamma(w)^q \right] = \int_{\mathbb{R}^q} \mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N, \epsilon}^\gamma(u_j) \right) \right] \prod_{k=1}^q w(u_k) du_k. \quad (2.12)$$

Then, Assumptions 2.2 and 2.4 imply that for almost all $\mathbf{u} \in \mathbb{A}^q$,

$$\lim_{\substack{N \rightarrow \infty \\ \epsilon \rightarrow 0}} \mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N, \epsilon}^\gamma(u_j) \right) \right] = \exp \left(\gamma^2 \sum_{1 \leq j < k \leq q} \mathbb{E} [G(u_j)G(u_k)] \right) \quad (2.13)$$

in both the weak and strong regime (as long as $\epsilon(N) \geq \delta_N$). Moreover, the condition (2.3) guarantees that for all $\mathbf{u} \in \mathbb{R}^q$,

$$\mathbb{E} \left[\exp \left(\sum_{k=1}^q \tilde{X}_{N, \epsilon}^\gamma(u_j) \right) \right] \leq C \prod_{1 \leq j < k \leq q} 1 \vee |u_j - u_k|^{-\gamma^2}.$$

Hence, if $\gamma^2 q < 2$, formula (2.8) follows directly from (2.12)–(2.13) and the dominated convergence theorem. \square

Note that in the context of Theorem 2.1 we did not require that the fields $X_{N,\epsilon}$, $X_{N+1,\epsilon}$, \dots are defined on the same probability space. However, if such a coupling is available, as in the case of the sine process, then we can upgrade the topology of convergence in Theorem 2.1 by replacing Lemma 2.2 by the following result.

Lemma 2.5. *Using the notation (1.27)–(1.29) where Λ_1 is the sine process. For any $w \in L^1 \cap L^\infty(\mathbb{R})$ and for any $q \geq 1$, the random variable $\mu_{N,\epsilon}^{\gamma, \text{Sine}}(w)$ converges in $L^q(\mathbb{P})$ as $N \rightarrow \infty$ to a limit $\mu_\epsilon^\gamma(w)$ whose law is the same as $v_\epsilon^\gamma(w)$.*

Before proving Lemma 2.5, let us first use the previous results to obtain Theorem 1.11.

Proof of Theorem 1.11. Let $\delta_N = N^{-1}(\log N)^{1+\kappa}$ where $\kappa > 0$, ϕ be a function which satisfies Assumption 1.1, and define for any $\mathbf{t} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^n$ and $\epsilon \in \Delta_n(\delta_N)$,

$$h_{\mathbf{u},\epsilon}(z) = \pi \sum_{k=1}^n t_k \int_{\mathbb{R}} \mathbb{1}_{|x-u_k| \leq \ell/2} \phi_{\epsilon_k}(z-x) dx.$$

This function is analytic in $|\Im z| \leq c\delta_N$ and we claim that it satisfies Assumption 1.2 in this strip with

$$C_\infty(h_{\mathbf{u},\epsilon}) = e^{\pi C_\phi |\mathbf{t}|} \quad \text{and} \quad C_1(h_{\mathbf{u},\epsilon}) = \pi C_\phi |\mathbf{t}|(\ell \vee 1),$$

where $|\mathbf{t}| = |t_1| + \dots + |t_n|$. To check this assumption, we can use the bounds:

$$\left| \int_{\mathbb{R}} \mathbb{1}_{|x-u| \leq \ell/2} \phi_\epsilon(z-x) dx \right| \leq \int_{\mathbb{R}} |\phi(x + i\Im z)| dx$$

and

$$\iint_{\mathbb{R} \times \mathbb{R}} \mathbb{1}_{|x-u| \leq \ell/2} |\phi_\epsilon(t-x+is)| dx dt \leq \ell \int_{\mathbb{R}} |\phi(x+is)| dx,$$

which hold for any $u \in \mathbb{R}$ and $\epsilon > 0$. This implies that we can apply Proposition 1.10. Moreover, since the sine process has constant density N on \mathbb{R} , the leading term in formula (1.32) corresponds to the expected value of the linear statistic $\sum_{\lambda \in \Lambda_N} h_{\mathbf{u},\epsilon}(\lambda)$, and by definition of the $H^{1/2}$ Gaussian noise Ξ (see Appendix A), the second order term corresponds to the variance of the Gaussian random variable $\Xi(h_{\mathbf{u},\epsilon})$. Thus, we get

$$\begin{aligned} \log \mathbb{E} \left[\exp \left(\sum_{\lambda \in \Lambda_N} h_{\mathbf{u},\epsilon}(\lambda) \right) \right] &= \mathbb{E} \left[\sum_{\lambda \in \Lambda_N} h_{\mathbf{u},\epsilon} \right] + \frac{\mathbb{E} [\Xi(h_{\mathbf{u},\epsilon})^2]}{2} \\ &\quad + O_{N \rightarrow \infty} \left(\delta_N^{-3/2} (|\mathbf{t}| \ell)^2 e^{\pi(6C_\phi |\mathbf{t}| - \delta_N N)} \right). \end{aligned}$$

By definition of the random field (1.27) and using the scaling property (1.28), we have the equality in law

$$\sum_{\lambda \in \Lambda_N} h_{\mathbf{u},\epsilon}(\lambda) \stackrel{d}{=} \sum_{k=1}^n t_k X_{N,\epsilon_k}(u_k),$$

and using the representation (A.14), we also have $\Xi(h_{\mathbf{u},\epsilon}) \stackrel{d}{=} \sum_{k=1}^n G_{\phi,\epsilon_k}(u_k)$. Hence, in the regime where $|\mathbf{t}|$ and $\ell > 0$ are independent of the parameter N , this implies that for any $\beta < 1 + \kappa$ and for any $\epsilon \in \Delta_n(\delta_N)$ with $\delta_N = (\log N)^{1+\kappa}/N$, we have

$$\begin{aligned} \log \mathbb{E} \left[\exp \left(\sum_{k=1}^n t_k \tilde{X}_{N,\epsilon_k}(u_k) \right) \right] &= \sum_{1 \leq k < j \leq n} t_k t_j \mathbb{E} [G_{\phi,\epsilon_j}(u_j) G_{\phi,\epsilon_k}(u_k)] \\ &+ O_{N \rightarrow \infty} \left(e^{-(\log N)^\beta} \right), \end{aligned} \quad (2.14)$$

uniformly for all $\mathbf{u} \in \mathbb{R}^n$. In particular, this immediately shows that the random field $u \mapsto X_{N,\epsilon}(u)$ satisfies Assumptions 2.1 and 2.4 for all $q \in \mathbb{N}$. Moreover, if the mollifier $\phi \in \mathcal{D}$, by Corollary 2.13, Assumption 2.2 holds too. Consequently, by Theorem 2.1, we obtain the convergence of the moments, formula (1.33). Then, if $w \in L^1 \cap L^\infty(\mathbb{R})$ and q is an even integer, according to Lemma 2.5, we have for any given $\epsilon > 0$, $\mu_{N,\epsilon}^{\gamma,\text{Sine}}(w)$ converges in $L^q(\mathbb{P})$ as $N \rightarrow \infty$ to the random variable $\mu_\epsilon^\gamma(w)$. In addition, if $\gamma^2 q < 2$, by Proposition B.1 and Remark B.2 below it, the random variable $\mu_\epsilon^\gamma(w)$ constructed above converges in $L^q(\mathbb{P})$ as $\epsilon \rightarrow 0$ to a random variable $\mu^\gamma(w)$ which has the same law as $v^\gamma(w)$. In other words, we have

$$\lim_{\epsilon \rightarrow 0, N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^{\gamma,\text{Sine}}(w) - \mu^\gamma(w) \right|^q \right] = 0.$$

On the other hand, by Lemma 2.4 and the triangle inequality, this implies that

$$\lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\delta_N}^{\gamma,\text{Sine}}(w) - \mu^\gamma(w) \right|^q \right] = 0$$

which concludes the proof of Theorem 1.9. \square

Proof of Lemma 2.5. To keep the notation simple, we will prove the Proposition when $q = 2$ which is the most interesting case. It is straightforward to generalize the argument to any even q . As in the proof of Theorem 1.9 (c.f. formula (2.14)), it is easy to check that the asymptotics of Proposition 1.10 implies that for any $\eta, \eta' \in \{0, 1\}$,

$$\log \mathbb{E} \left[\exp \left(\tilde{X}_{N+\eta,\epsilon}^\gamma(u) + \tilde{X}_{N+\eta',\epsilon}^\gamma(v) \right) \right] = \gamma^2 T_{\epsilon,\epsilon}(u, v) + O_{\eta,\eta'} \left(e^{-(\log N)^\beta} \right)$$

uniformly for all $u, v \in \mathbb{R}$, where $T_{\epsilon,\epsilon}(u, v) = \mathbb{E} [G_{\phi,\epsilon}(u) G_{\phi,\epsilon}(v)]$ and $\beta < 1 + \kappa$. By expanding the square, this implies that

$$\begin{aligned} &\mathbb{E} \left[\left| \mu_{N,\epsilon}^{\gamma,\text{Sine}}(w) - \mu_{N+1,\epsilon}^{\gamma,\text{Sine}}(w) \right|^2 \right] \\ &= \int_{\mathbb{R}} e^{\gamma^2 T_{\epsilon,\epsilon}(u,v)} \left\{ \sum_{\eta,\eta' \in \{0,1\}} (-1)^{\eta+\eta'} \exp O_{\eta,\eta'} \left(e^{-(\log N)^\beta} \right) \right\} w(u) w(v) du dv. \end{aligned} \quad (2.15)$$

Note that the leading terms on the RHS of formula (2.15) cancel and the error terms are uniform. Moreover, by (2.3), $T_{\epsilon,\epsilon}(u, v) \leq \log^+(\epsilon^{-1}) + C$ on \mathbb{R}^2 and since $w \in L^1(\mathbb{R})$, we obtain for any $\epsilon > 0$,

$$\mathbb{E} \left[\left| \mu_{N,\epsilon}^{\gamma,\text{Sine}}(w) - \mu_{N+1,\epsilon}^{\gamma,\text{Sine}}(w) \right|^2 \right] = O_{w,\epsilon} \left(e^{-(\log N)^\beta} \right).$$

Thus, since we may choose the parameter $\beta > 1$, by the triangle inequality, this shows that $(\mu_{N,\epsilon}^{\gamma,\text{Sine}}(w))_{N>0}$ is a Cauchy sequence in $L^2(\mathbb{P})$. Let us denote by $\mu_\epsilon^\gamma(w)$ its limit. We will use Lemma 2.2 to identify the law of this limit. For any $M > 0$, we let $w_M(u) = w(u)\mathbb{1}_{\{|u|\leq M\}}$ and $w_M^*(u) = w(u)\mathbb{1}_{\{|u|>M\}}$. On the one hand, by formula (2.12) and using Assumption 2.4 and the bound (2.3), we obtain

$$\lim_{N\rightarrow\infty} \mathbb{E} \left[\mu_{N,\epsilon}^{\gamma,\text{Sine}}(w_M^*)^2 \right] = \mathbb{E} \left[\mu_\epsilon^\gamma(w_M^*)^2 \right] \ll \epsilon^{-1} \int_{\substack{|u|>M \\ |v|>M}} w(u)w(v)dudv.$$

In addition, we easily check that $\mu_\epsilon^\gamma(w) = \mu_\epsilon^\gamma(w_M) + \mu_\epsilon^\gamma(w_M^*)$ so that by taking the limit as $M \rightarrow \infty$,

$$\mu_\epsilon^\gamma(w_M) \Rightarrow \mu_\epsilon^\gamma(w).$$

An analogous computation shows that for the regularized GMC measure: $v_\epsilon^\gamma(w_M) \Rightarrow v_\epsilon^\gamma(w)$ as $M \rightarrow \infty$. On the other hand, since the functions w_M have compact support, by Lemma 2.2, we know that $\mu_\epsilon^\gamma(w_M) \stackrel{d}{=} v_\epsilon^\gamma(w_M)$ for any $M > 0$. Hence, we conclude that $\mu_\epsilon^\gamma(w) \stackrel{d}{=} v_\epsilon^\gamma(w)$. \square

2.2. Convergence of the multiplicative chaos measure in the L^1 -phase. We work in the same context as in the previous section and the goal is to establish the convergence of the random measure $\mu_{N,\epsilon}^\gamma(du) = \exp(\tilde{X}_{N,\epsilon}^\gamma(u))du$ throughout the subcritical phase $0 < \gamma < \sqrt{2d}$.

Theorem 2.6. *Suppose that Assumptions 2.1–2.5 hold for $q \leq 3$ in 2.4. If $w \in L^1 \cap L^\infty(\mathfrak{A})$ and $\gamma < \sqrt{2d}$, then for any sequence $\epsilon_N \rightarrow 0$ in such a way that $\epsilon_N \geq \delta_N$, the random variable $\mu_{N,\epsilon_N}^\gamma(w)$ converges in distribution to $v^\gamma(w)$ as $N \rightarrow \infty$.*

The condition $\gamma < \sqrt{2d}$ is sharp in the sense that we expect that $\mu_{N,\epsilon}^\gamma(du) \rightarrow 0$ for any $\gamma \geq \sqrt{2d}$. The proof of Theorem 2.6 follows the elementary argument introduced [5] to prove convergence of the GMC measure in the L^1 -phase. In fact, the asymptotics (2.7) are so strong that Berestycki's method can be applied to the field $u \mapsto X_{N,\epsilon}(u)$ modulo a few technical issues. The main idea stems from the fact that the measure $\mu_{N,\epsilon}^\gamma$ is supported on the so-called γ -thick points, (1.18). More specifically, we will proceed to show that, if the parameter $\alpha > \gamma$, the mass

$$\mu_{N,\epsilon_N}^\gamma \left(u \in \mathfrak{A} : \frac{X_{N,\epsilon}(u)}{\log \epsilon^{-1}} > \alpha \text{ for some } \epsilon \in \{e^{-k} : L \leq k \leq \log \delta_N^{-1}\} \right)$$

converges to 0 in $L^1(\mathbb{P})$ as $N \rightarrow \infty$ and then $L \rightarrow \infty$. Then we will show that the random measures $\mu_{N,\epsilon_N}^\gamma$ restricted to the good set $\left\{ u \in \mathfrak{A} : \frac{X_{N,\epsilon}(u)}{\log \epsilon^{-1}} \leq \alpha \text{ for all } \epsilon \in \{e^{-k} : L \leq k \leq \log \delta_N^{-1}\} \right\}$ form a Cauchy net in $L^1(\mathbb{P})$ (in the sense of Proposition 2.9) if α is sufficiently close to γ and $\gamma < \sqrt{2d}$. Like in Sect. 2.1, we will rely on the fact that the weak and strong limits coincide to identify that the law of the random variable μ^γ is the same law as the GMC measure v^γ . First of all, we need to introduce further notation and establish some preparatory lemmas. Then, we will give the proofs of Theorems 2.6, 1.7 and finally of Theorem 1.9.

For any $u \in \mathfrak{A}$ and $N > 0$, let $Z^N(u)$ be the random variable taking values in \mathbb{R}^N (measurable with respect to the process $(X_{N,\epsilon})_{\epsilon>0}$) given by

$$Z_k^N(u) = X_{N,e^{-k}}(u). \quad (2.16)$$

Here, \mathbb{R}^N is equipped with the usual product topology and its Borel σ -algebra. For any $u \in \mathfrak{A}$, let $\mathbb{P}_{N,\epsilon}^u$ be the probability measure with Radon–Nykodym derivative proportional to $\exp(\tilde{X}_{N,\epsilon}^u(u))$ with respect to the probability measure \mathbb{P} . Finally, for any $(u, v) \in \mathfrak{A}^2$ such that $u \neq v$, let $\mathbb{P}_{N,\epsilon}^{u,v}$ be the probability measure with Radon–Nykodym derivative proportional to $\exp(\tilde{X}_{N,\epsilon}^u(u) + \tilde{X}_{N,\epsilon}^v(v))$ and, in the mixed regime, let $\tilde{\mathbb{P}}_{N,\epsilon}^{u,v}$ be the probability measure with Radon–Nykodym derivative proportional to $\exp(\tilde{X}_{N,\delta_N}^u(u) + \tilde{X}_{N,\epsilon}^v(v))$ with respect to \mathbb{P} . We make the following assumption:

Assumption 2.5 (*Weak convergence of finite-dimensional distributions under the biased measures*). For any $(u, v) \in \mathfrak{A}^2$ such that $u \neq v$, one has

$$\begin{aligned} \text{Law}_{\mathbb{P}_{N,\epsilon_N}^{u,v}}(Z^N(u), Z^N(v)) &\rightarrow \mathbb{G}_{u,v} \quad \text{as } N \rightarrow \infty, \\ \text{Law}_{\mathbb{P}_{N,\epsilon}^{u,v}}(Z^N(u), Z^N(v)) &\rightarrow \mathbb{G}_{u,v} \quad \text{as } N \rightarrow \infty \text{ and then } \epsilon \rightarrow 0, \\ \text{Law}_{\tilde{\mathbb{P}}_{N,\epsilon}^{u,v}}(Z^N(u), Z^N(v)) &\rightarrow \mathbb{G}_{u,v} \quad \text{as } N \rightarrow \infty \text{ and then } \epsilon \rightarrow 0, \end{aligned}$$

where the convergence holds weakly for finite dimensional marginals and $\mathbb{G}_{u,v}$ is a Gaussian measure on $\mathbb{R}^N \times \mathbb{R}^N$ with mean

$$\mathbf{E}_{\mathbb{G}_{u,v}}[Z_k(x)] = \gamma(T_{0,e^{-k}}(x, x) + T_{0,e^{-k}}(u, v)), \quad \forall x \in \{u, v\}, \quad (2.17)$$

and covariance structure:

$$\langle Z_k(x); Z_j(y) \rangle_{\mathbb{G}_{u,v}} = T_{e^{-k}, e^{-j}}(x, y), \quad \forall x, y \in \{u, v\}. \quad (2.18)$$

For any $\alpha > 0$ and $0 < L < M$, we define the following events:

$$A_{L,M}^\alpha(Z) = \bigcap_{k=L}^M \{Z_k < \alpha k\} \quad \text{and} \quad A_{L,M}^{\alpha*}(Z) = \{\exists k \in [L, M] : Z_k \geq \alpha k\}. \quad (2.19)$$

Observe that one has $A_{L,\infty}^\alpha(Z) = \bigcap_{M=L}^\infty A_{L,M}^\alpha(Z)$ and for any $0 < L < L' < M$,

$$A_{L,L'}^\alpha(Z) \cap A_{L,M}^{\alpha*}(Z) = A_{L,L'}^\alpha(Z) \cap A_{L',M}^{\alpha*}(Z). \quad (2.20)$$

Lemma 2.7. *Let $M_N = \lfloor \log \delta_N^{-1} \rfloor$. If $\gamma < \alpha$ and L is sufficiently large, one has for any $\epsilon > 0$,*

$$\limsup_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon}^u \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] \ll e^{-\frac{(\alpha-\gamma)^2}{4}L}.$$

Proof. By a union bound,

$$\mathbb{P}_{N,\epsilon}^u \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] \leq \sum_{k=L}^{\infty} \mathbb{P}_{N,\epsilon}^u \left[Z_k^N(u) \geq \alpha k \right].$$

By Assumption 2.1, under the law $\mathbb{P}_{N,\epsilon}^\mu$, $Z_k^N(u)$ converges weakly to Gaussian random variable with mean $\gamma T_{\epsilon,e^{-k}}(u, u)$ and variance $T_{e^{-k},e^{-k}}(u, u)$ for any $k \in \mathbb{N}$. Therefore, by the Portmanteau Theorem and a Gaussian tail-bound, we obtain

$$\limsup_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon}^\mu \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] \leq \sum_{k=L}^{\infty} \exp \left(- \frac{(\alpha k - \gamma T_{\epsilon,e^{-k}}(u, u))^2}{2 T_{e^{-k},e^{-k}}(u, u)} \right).$$

Thus, by Assumption 2.2, if $\alpha > \gamma$ and the parameter L is sufficiently large, we can assume that for all $k \geq L$,

$$\alpha k - \gamma T_{\epsilon,e^{-k}}(u, u) \geq \sqrt{3}(\alpha - \gamma)k/2 \quad \text{and} \quad T_{e^{-k},e^{-k}}(u, u) \leq 3k/2.$$

This implies that for any $u \in \mathfrak{A}$,

$$\limsup_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon}^\mu \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] \leq \sum_{k \geq L} \exp \left(- \frac{(\alpha - \gamma)^2}{4} k \right)$$

which completes the proof. \square

Lemma 2.8. *Let $M_N = \lfloor \log \delta_N^{-1} \rfloor$ and suppose that $\gamma < \alpha < 2\gamma$. One has for any $u \in \mathfrak{A}$,*

$$\lim_{L \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon_N}^\mu \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] = 0.$$

Similarly, one has for any points $u, v \in \mathfrak{A}$ and $x \in \{u, v\}$,

$$\lim_{L \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^{\alpha*}(Z^N(x)) \right] = 0.$$

Proof. By a union bound and Markov's inequality, we obtain for any $t > 0$,

$$\begin{aligned} \mathbb{P}_{N,\epsilon_N}^\mu \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] &\leq \sum_{k=L}^{M_N} \mathbb{P}_{N,\epsilon_N}^\mu \left[Z_k^N(u) \geq \alpha k \right] \\ &\leq \sum_{k=L}^{M_N} e^{-t\alpha k + \frac{t^2}{2} \mathbb{E}[X_{N,e^{-k}}^2(u)]} \mathbb{E} \left[e^{\tilde{X}_{N,e^{-k}}^t(u) + \tilde{X}_{N,\epsilon_N}^\gamma(u)} \right] \end{aligned}$$

Using the assumptions, in particular the asymptotics (2.7), this implies that

$$\mathbb{P}_{N,\epsilon_N}^\mu \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] \ll \sum_{k=L}^{M_N} e^{-t\alpha k + \frac{t^2}{2} T_{e^{-k},e^{-k}}(u,u) + t\gamma T_{e^{-k},\epsilon_N}(u,u)}.$$

Using the upper-bound (2.3) and choosing $t = \alpha - \gamma$, this shows that

$$\begin{aligned} \limsup_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon_N}^\mu \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] &\ll \sum_{k=L}^{\infty} e^{-tk((\alpha-\gamma)-t/2)} \\ &\ll e^{-L(\alpha-\gamma)^2/2}. \end{aligned}$$

Letting $L \rightarrow \infty$, this completes the first part of the proof. Now, for the second estimate, an analogous argument shows that

$$\begin{aligned} & \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^{\alpha*}(Z^N(x)) \right] \\ & \ll \sum_{k=L}^{\infty} e^{-t\alpha k + \frac{t^2}{2} T_{e^{-k},e^{-k}}(u,u) + t\gamma (T_{e^{-k},\epsilon_N}(x,u) + T_{e^{-k},\epsilon_N}(x,v)) + \gamma^2 T_{\epsilon_N,\epsilon_N}(u,v)}. \end{aligned}$$

Then, if $L \geq \zeta = \log |u - v|^{-1}$, still choosing $t = \alpha - \gamma < \gamma$, we obtain

$$\begin{aligned} \limsup_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^{\alpha*}(Z^N(x)) \right] & \ll e^{2\gamma^2 \zeta} \sum_{k=L}^{\infty} e^{-tk((\alpha-\gamma)-t/2)} \\ & \ll e^{-L(\alpha-\gamma)^2/2+2\gamma^2\zeta} \end{aligned}$$

which converges to 0 as $L \rightarrow \infty$. \square

Proposition 2.9. *Let $M_N = \lfloor \log \delta_N^{-1} \rfloor$ and $0 < \gamma < \sqrt{2d}$. There exists $\gamma < \alpha < 2\gamma$ so that, under the assumptions of Theorem 2.6, one has*

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma(w \mathbb{1}_{A_{L,M_N}^\alpha(Z^N)}) - \mu_{N,\epsilon}^\gamma(w \mathbb{1}_{A_{L,M_N}^\alpha(Z^N)}) \right|^2 \right] = 0.$$

Proof. In this proof, the parameter L is assumed to be large but fixed. For any $(u, v) \in \mathfrak{A}^2$, $\epsilon > 0$ and $N > 0$, let us define

$$\mathcal{W}_{N,\epsilon}^{\mu,v} = \mathbb{E} \left[\exp \left(\tilde{X}_{N,\epsilon}(u) + \tilde{X}_{N,\epsilon}(v) \right) \right] \quad \text{and} \quad \tilde{\mathcal{W}}_{N,\epsilon}^{\mu,v} = \mathbb{E} \left[\exp \left(\tilde{X}_{N,\epsilon_N}(u) + \tilde{X}_{N,\epsilon}(v) \right) \right].$$

We also let

$$\mathcal{I}_{N,\epsilon} = \iint_{\mathfrak{A}^2} \mathbb{P}_{N,\epsilon}^{\mu,v} [A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v))] \mathcal{W}_{N,\epsilon}^{\mu,v} w(u) w(v) du dv$$

and

$$\tilde{\mathcal{I}}_{N,\epsilon} = \iint_{\mathfrak{A}^2} \tilde{\mathbb{P}}_{N,\epsilon}^{\mu,v} [A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v))] \tilde{\mathcal{W}}_{N,\epsilon}^{\mu,v} w(u) w(v) du dv.$$

As for the proof of Proposition 2.4, we may expand

$$\mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma(w \mathbb{1}_{A_{L,M_N}^\alpha(Z^N)}) - \mu_{N,\epsilon}^\gamma(w \mathbb{1}_{A_{L,M_N}^\alpha(Z^N)}) \right|^2 \right] = \mathcal{I}_{N,\epsilon_N} + \mathcal{I}_{N,\epsilon} - 2\tilde{\mathcal{I}}_{N,\epsilon} \quad (2.21)$$

and we would like to prove that all the terms converge to the same limit when $N \rightarrow \infty$ and then $\epsilon \rightarrow 0$. We will focus on computing the limit of the integral $\mathcal{I}_{N,\epsilon_N}$ as $N \rightarrow \infty$. On the one hand by (2.20) and elementary algebra, we obtain for any $L < L' < M_N$,

$$\begin{aligned} \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v)) \right] &= \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,L'}^\alpha(Z^N(u)) \cap A_{L,L'}^\alpha(Z^N(v)) \right] \\ &- \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,L'}^\alpha(Z^N(u)) \cap A_{L,L'}^\alpha(Z^N(v)) \cap A_{L',M_N}^{\alpha*}(Z^N(u)) \right] \\ &- \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,L'}^\alpha(Z^N(v)) \cap A_{L',M_N}^{\alpha*}(Z^N(v)) \cap A_{L,M_N}^\alpha(Z^N(u)) \right]. \end{aligned}$$

The Assumption 2.5 implies that, under the law $\mathbb{P}_{N,\epsilon_N}^{\mu,v}$, $(Z_k^N(u), Z_k^N(v))_{k=L}^{L'}$ converges weakly to a certain Gaussian vector. Then, by Lemma 2.8, this implies that for any $u \neq v$,

$$\begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v)) \right] \\ &= \mathbb{G}_{u,v} \left[A_{L,L'}^\alpha(Z(u)) \cap A_{L,L'}^\alpha(Z(v)) \right] + o(1)_{L' \rightarrow \infty}. \end{aligned}$$

Consequently, by the monotone convergence theorem, we obtain

$$\begin{aligned} & \lim_{N \rightarrow \infty} \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v)) \right] \\ &= \mathbb{G}_{u,v} \left[A_{L,\infty}^\alpha(Z(u)) \cap A_{L,\infty}^\alpha(Z(v)) \right]. \end{aligned} \quad (2.22)$$

On the other hand, Assumption 2.4 implies that for all $(u, v) \in \mathfrak{A}^2$,

$$\mathcal{W}_{N,\epsilon_N}^{u,v} \sim \exp(\gamma^2 T_{\epsilon_N,\epsilon_N}(u, v)) \quad \text{as } N \rightarrow \infty. \quad (2.23)$$

So, in the regime where $\gamma > \sqrt{d}$, we expect the limit of the integral $\mathcal{I}_{N,\epsilon_N}$ to be finite only if the probability $\mathbb{G}_{u,v} \left[A_{L,\infty}^\alpha(Z(u)), A_{L,\infty}^\alpha(Z(v)) \right]$ converges to 0 sufficiently fast as $|u - v| \rightarrow 0$. In order to prove this, we use that $A_{L,M_N}^\alpha(Z) \subset \{Z_\zeta \leq \alpha\zeta\}$ where we choose

$$\zeta = \begin{cases} \log |u - v|^{-1} & \text{if } L \leq \log |u - v|^{-1} \leq \log \epsilon_N^{-1} \\ L & \text{if } \log |u - v|^{-1} \leq L \\ \log \epsilon_N^{-1} & \text{if } |u - v| \leq \epsilon_N \end{cases}$$

so that

$$\mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v)) \right] \leq \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[Z_\zeta^N(u) \leq \alpha\zeta \right].$$

By Markov's inequality and the asymptotics (2.7), this implies that for any $t > 0$,

$$\begin{aligned} & \mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[Z_\zeta^N(u) \leq \alpha\zeta \right] \\ & \leq \frac{\mathbb{E} \left[\exp(\gamma X_{N,\epsilon_N}(u) + \gamma X_{N,\epsilon_N}(v) - t X_{N,e^{-\zeta}}(u)) \right]}{\mathbb{E} \left[\exp(\gamma X_{N,\epsilon_N}(u) + \gamma X_{N,\epsilon_N}(v)) \right]} e^{\alpha t \zeta} \\ & = \exp \left(\alpha t \zeta - \gamma t (T_{\epsilon_N,e^{-\zeta}}(u, u) + T_{\epsilon_N,e^{-\zeta}}(u, v)) + \frac{t^2}{2} T_{e^{-\zeta},e^{-\zeta}}(u, u) + o(1)_{N \rightarrow \infty} \right). \end{aligned}$$

Then, choosing $t = t_*(N, \gamma, \alpha, u, v) := \frac{\gamma T_{\epsilon_N,e^{-\zeta}}(u,u) + \gamma T_{\epsilon_N,e^{-\zeta}}(u,v) - \alpha\zeta}{T_{e^{-\zeta},e^{-\zeta}}(u,u)}$, we have

$$\mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[Z_\zeta^N \leq \alpha\zeta \right] \ll \exp \left(-\frac{t_*^2}{2} T_{e^{-\zeta},e^{-\zeta}}(u, u) \right). \quad (2.24)$$

Note that by definition of ζ , Assumption 2.2 implies that $T_{\epsilon_N,e^{-\zeta}}(u, v) = \zeta + O(1)_{L \rightarrow \infty}$ when $|u - v| \leq e^{-L}$. In particular, in the regime $|u - v| \leq e^{-L}$, by (2.5), we see that

$t_* = (2\gamma - \alpha) + o(1)$. Hence, if the parameter L is sufficiently large and $\alpha < 2\gamma$, the parameter $t^* > 0$ so that the bound (2.24) holds and we obtain

$$\mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[Z_\zeta^N \leq \alpha \zeta \right] \ll \exp \left(-\frac{(2\gamma - \alpha)^2}{2} \zeta \right).$$

This shows that for all $u, v \in \mathfrak{A}$ such that $|u - v| \leq e^{-L}$,

$$\mathbb{P}_{N,\epsilon_N}^{\mu,v} \left[A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v)) \right] \ll (|u - v| \vee \epsilon_N)^{\frac{(2\gamma - \alpha)^2}{2}}.$$

Moreover, in the regime $|u - v| > e^{-L}$, the RHS of (2.23) remains bounded as $N \rightarrow \infty$. Hence, using the bound (2.3), these estimates show that

$$\begin{aligned} \mathbb{P}_{N,\epsilon}^{\mu,v} [A_{L,M_N}^\alpha(Z^N(u)) \cap A_{L,M_N}^\alpha(Z^N(v))] \mathscr{W}_{N,\epsilon_N}^{\mu,v} &\ll (|u - v| \vee \epsilon_N)^{\frac{(2\gamma - \alpha)^2}{2} - \gamma^2} \\ &\ll |u - v|^{-\gamma^2 + \frac{(2\gamma - \alpha)^2}{2}}. \end{aligned} \quad (2.25)$$

The LHS of (2.25) is locally integrable on $\mathbb{R}^d \times \mathbb{R}^d$ if $\gamma^2 - \frac{(2\gamma - \alpha)^2}{2} < d$. Hence, as long as $\gamma^2 < 2d$, it is possible to choose $\alpha > \gamma$ so that this condition is satisfied. By the dominated convergence theorem, formulae (2.22)–(2.23) and (2.4), we conclude that $\lim_{N \rightarrow \infty} \mathcal{I}_{N,\epsilon_N}^{\gamma,L} = \mathcal{I}_\infty^{\gamma,L}$ where

$$\mathcal{I}_\infty^{\gamma,L} := \iint_{\mathbb{R}^2} \mathbb{G}_{u,v} [A_{L,\infty}^\alpha(Z(u)) \cap A_{L,\infty}^\alpha(Z(v))] e^{\gamma^2 T(u,v)} w(u) w(v) du dv < \infty.$$

We can apply the same argument in the weak and mixed regimes as well and obtain

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathcal{I}_{N,\epsilon} = \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \tilde{\mathcal{I}}_{N,\epsilon} = \mathcal{I}_\infty^{\gamma,L}.$$

In the end, combining these limits with formula (2.21), we have completed the proof. \square

Proof of Theorem 2.6. Recall that by Lemma 2.2, for any $\epsilon > 0$, $\mu_{N,\epsilon}^\gamma(w) \Rightarrow v_\epsilon^\gamma(w)$ as $N \rightarrow \infty$ and, according to Assumption 2.3, if $\gamma < \sqrt{2d}$, the random variable $v_\epsilon^\gamma(w) \Rightarrow v^\gamma(w)$ as $\epsilon \rightarrow 0$. Therefore, by [43, Theorem 4.28], it suffices to establish that

$$\limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right| \right] = 0.$$

By the triangle inequality, we have for any $L, \alpha > 0$ and $\epsilon > 0$,

$$\begin{aligned} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right| \right] &\leq \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^\alpha(Z^N)} \right) - \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^\alpha(Z^N)} \right) \right| \right] \\ &\quad + \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right] \\ &\quad + \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right]. \end{aligned}$$

Then, by Proposition 2.9, if $\gamma < \sqrt{2d}$, by choosing the parameter $\alpha > \gamma$ appropriately, we obtain

$$\begin{aligned} \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma(w) - \mu_{N,\epsilon}^\gamma(w) \right| \right] &\leq \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right] \\ &+ \limsup_{\epsilon \rightarrow 0} \limsup_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right]. \end{aligned} \quad (2.26)$$

Since the LHS of (2.26) does not depend on the parameter $L > 0$, in order to complete the proof, it suffices to show that

$$\lim_{L \rightarrow \infty} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right] = 0, \quad (2.27)$$

$$\lim_{L \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right] = 0. \quad (2.28)$$

By definition of the probability measure $\mathbb{P}_{N,\epsilon}^u$, one has for any $\epsilon > 0$,

$$\mathbb{E} \left[\left| \mu_{N,\epsilon}^\gamma \left(w \mathbb{1}_{A_{L,M_N}^{\alpha*}(Z^N)} \right) \right| \right] \leq \int_{\mathfrak{A}} \mathbb{P}_{N,\epsilon}^u \left[A_{L,M_N}^{\alpha*}(Z^N(u)) \right] \mathbb{E} \left[e^{\tilde{X}_{N,\epsilon}(u)} \right] |w(u)| du.$$

Note that by Assumption 2.4, $\mathbb{E} \left[e^{\tilde{X}_{N,\epsilon}(u)} \right] \rightarrow 1$ as $N \rightarrow \infty$ and $\epsilon \rightarrow 0$ uniformly for all $u \in \mathfrak{A}$, so that by Lemma 2.7, one obtains (2.28). Finally, (2.27) follows from an analogous argument using the estimate of Lemma 2.8. \square

In principle, it is possible to verify Assumption 2.5 without knowing the asymptotics of all exponential moments of the random field $X_{N,\epsilon}(u)$. However, if the Assumption 2.4 holds for all $q \in \mathbb{N}$, then the Assumption 2.5 follows from a routine computation. In particular, as a consequence of the asymptotics obtained in Sects. 2.3 and 3, this applies to the CUE and sine process.

Proof of Theorem 1.7. It suffices to check that the fields in question satisfy the assumptions of Theorem 2.6. First of all, the strong Gaussian asymptotics (1.20) implies directly Assumption 2.4 as well as the fact that, for fixed $\epsilon > 0$, the field $u \mapsto X_{N,\epsilon}(u)$ converges in the sense of finite-dimensional distributions to a mean-zero Gaussian process G_ϵ . In addition, one has for any $\mathbf{t} \in \mathbb{R}^q$ and $\mathbf{u} \in \{u, v\}^q$,

$$\begin{aligned} &\mathbf{E}_{\mathbb{P}_{N,\epsilon_N}^{\mu,v}} \left[\exp \left(\sum_{k=1}^q t_k X_{N,e^{-k}}(u_k) \right) \right] \\ &= \frac{\mathbb{E} \left[\exp \left(\gamma X_{N,\delta_N}(u) + \gamma X_{N,\delta_N}(v) + \sum_{k=1}^q t_k X_{N,e^{-k}}(u_k) \right) \right]}{\mathbb{E} \left[\exp \left(\gamma X_{N,\delta_N}(u) + \gamma X_{N,\delta_N}(v) \right) \right]} \\ &= \exp \left(\sum_{k=1}^q t_k \mathbf{E}_{G_{u,v}}[Z_k(u_k)] + \frac{1}{2} \sum_{k,j=1}^q t_k t_j \langle Z_k(u_k); Z_j(u_j) \rangle_{G_{u,v}} + o(1) \right)_{N \rightarrow \infty} \end{aligned}$$

according to (2.17)–(2.18). This establishes, the first part of the Assumption 2.5; the other assertions in the weak and mixed regimes are checked in a similar fashion. Finally, when $G_\epsilon = G * \phi_\epsilon$, we check in Sect. 2.3 that the covariance kernel satisfies both Assumptions 2.1 and 2.2 (the computation is given for the stationary field G with correlation kernel (1.22) but they can be easily generalized in other situations—see for instance [5]). In this context, the condition 2.3 (which, in general, is a non-trivial issue) follows from Theorem 1.2 presented in the introduction. \square

Proof of Theorem 1.9. If $\phi \in \mathcal{D}$ is a mollifier which satisfies Assumption 1.1, formula (2.14) shows that the random field $X_{N,\epsilon}$ given by (1.27) satisfies the strong Gaussian asymptotics of Assumption 2.4 with $G_\epsilon = G * \phi_\epsilon$. In this context, by Theorem 1.2, we know that for any $w \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ and $\gamma < \sqrt{2d}$, the random variable $v_\epsilon^\gamma(w)$ converges in $L^1(\mathbb{P})$ to $v^\gamma(w)$ as $\epsilon \rightarrow 0$. In addition, we have already checked (see the proof of Theorem 1.7 above) that the Assumptions 2.1–2.5 are also satisfied. Then, by Lemma 2.5, for any $\epsilon > 0$, $\mu_{N,\epsilon}^{\gamma, \text{Sine}}(w)$ converges in $L^1(\mathbb{P})$ to some random variable $\mu_\epsilon^\gamma(w)$ and, since $\mu_\epsilon^\gamma(w) \stackrel{d}{=} v_\epsilon^\gamma(w)$, this shows that for any $\gamma < \sqrt{2d}$, there exists a random variable $\mu^\gamma(w) \stackrel{d}{=} v^\gamma(w)$ so that by taking successive limits, we obtain

$$\lim_{\epsilon \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon}^{\gamma, \text{Sine}}(w) - \mu^\gamma(w) \right| \right] = 0. \quad (2.29)$$

Then, using Lemmas 2.7, 2.8 and Proposition 2.9, proceeding exactly as in the proof of Theorem 2.6, we can show that if $\gamma < \sqrt{2d}$ and $\epsilon_N \geq \delta_N N (\log N)^{1+\kappa}$ for some $\kappa > 0$, we have

$$\lim_{\epsilon_N \rightarrow 0} \lim_{N \rightarrow \infty} \mathbb{E} \left[\left| \mu_{N,\epsilon_N}^{\gamma, \text{Sine}}(w) - \mu_{N,\epsilon}^{\gamma, \text{Sine}}(w) \right| \right] = 0. \quad (2.30)$$

Using the triangle inequality and combining the limits (2.29) and (2.30), we conclude that the random variable $\mu_{N,\epsilon_N}^{\gamma, \text{Sine}}(w)$ converges in $L^1(\mathbb{P})$ to $\mu^\gamma(w)$ as $N \rightarrow \infty$ and $\mu^\gamma(w) \stackrel{d}{=} v^\gamma(w)$. \square

2.3. Asymptotics of the covariances. Let G be the stationary Gaussian process on \mathbb{R} with covariance function Q , (1.22), and recall that, for any mollifier ϕ , we denote $G_{\phi,\epsilon} = G * \phi_\epsilon$. In this section, we derive the asymptotics of the covariance

$$\mathbb{E} [G_{\phi,\epsilon}(u) G_{\psi,\delta}(v)] = \int_{\mathbb{R}} e^{-2\pi i(u-v)\kappa} \hat{\psi}(\epsilon\kappa) \overline{\hat{\phi}(\delta\kappa)} \widehat{Q}(\kappa) d\kappa, \quad (2.31)$$

as $\epsilon, \delta \rightarrow 0$, for all $u, v \in \mathbb{R}$ and for a large class of mollifiers. This is relevant to check that the conditions (2.4), (2.3), and (2.5) are satisfied and apply Theorems 2.1 and 2.6 to conclude that the multiplicative chaos measures associated with counting statistics of the CUE and sine process exist and have the same law. For any $\ell, \epsilon > 0$, we define the function

$$Q_\epsilon(x) = -\log \left(\frac{\epsilon}{2\pi} \vee |x| \right) + \log \left(\frac{\epsilon}{2\pi} \vee \sqrt{|\ell^2 - x^2|} \right). \quad (2.32)$$

In the following, we will use the notation $\mathcal{U}_{\mathbf{u}}^{\epsilon \rightarrow 0}$ to specify a function of the variable \mathbf{u} and the parameter $\epsilon > 0$ which is uniformly bounded (by a universal constant) and converges to 0 as $\epsilon \rightarrow 0$ for almost all $\mathbf{u} \in \mathbb{R}^d$.

Recall that we defined $\mathcal{D} = \bigcup_{\alpha>0} \mathcal{D}_\alpha$, c.f. (1.30), and recall also the definition of the cosine integrals:

$$\text{Cin}(\omega) = \int_0^\omega \frac{1 - \cos(z)}{z} dz \quad \text{and} \quad \text{Ci}(x) = \int_1^\infty \frac{\cos(xt)}{t} dt.$$

The function Cin is entire, while the function Ci is even on \mathbb{R} with the value $+\infty$ at 0, and it turns out that for any $\omega \in \mathbb{R} \setminus \{0\}$,

$$\text{Cin}(\omega) = \text{Ci}(\omega) + \log |\omega| + \gamma,$$

where γ is the Euler constant. In particular, since $\lim_{x \rightarrow \infty} \text{Ci}(x) = 0$, we have for any $\omega \in \mathbb{R}$,

$$\text{Cin}(\omega/\epsilon) = \log^+(\omega/\epsilon) + \gamma + \mathcal{U}_\omega. \quad (2.33)$$

Lemma 2.10. *Let $\Phi \in L^1(\mathbb{R}_+)$, continuous with $\Phi(0) = 1$, and so that the function $\kappa \mapsto \frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa}$ is integrable on \mathbb{R}_+ . Then, the function*

$$\mathcal{E}_\Phi(\omega) = \int_0^\infty (1 - \cos(\omega\kappa)) \frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa} d\kappa$$

is continuously differentiable on \mathbb{R} , and $\lim_{\omega \rightarrow \infty} \mathcal{E}_\Phi(\omega) = \int_0^\infty \frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa} d\kappa$.

Moreover, for all $\omega \in \mathbb{R}$,

$$\int_0^\infty \frac{1 - \cos(\omega\kappa)}{\kappa} \Phi(\kappa) d\kappa = \text{Cin}(\omega) + \mathcal{E}_\Phi(\omega). \quad (2.34)$$

Proof. All the properties of the function \mathcal{E}_Φ are easy to check. In particular, we have

$$\mathcal{E}'_\Phi(\omega) = \int_0^\infty \sin(\omega\kappa) (\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}) d\kappa,$$

and the limit of $\mathcal{E}_\Phi(\omega)$ as $\omega \rightarrow \infty$ follows directly from the Riemann–Lebesgue lemma. Finally, the identity (2.34) is an immediate consequence of the definition of the cosine integral. \square

Proposition 2.11. *For any functions $\phi, \psi \in \mathcal{D}$, we have for all $u, v \in \mathbb{R}$,*

$$\mathbb{E} [G_{\phi, \epsilon}(u) G_{\psi, \epsilon}(v)] = Q_\epsilon(u - v) + \mathcal{U}_{u, v}. \quad \epsilon \rightarrow 0$$

Proof. Let $\Phi = \Re\{\widehat{\phi\psi}\}$, $\Psi = \Im\{\widehat{\phi\psi}\}$, and $\epsilon' = \epsilon/2\pi$. We claim that the function Φ satisfies the assumption of Lemma 2.10. Namely, by the Cauchy–Schwarz inequality, $\Phi \in L^1(\mathbb{R})$ and by Plancherel’s formula, for any $\kappa \in \mathbb{R}$,

$$\Phi(\kappa) = \iint_{\mathbb{R}^2} \cos(2\pi\kappa x) \phi(x+t) \psi(t) dx dt.$$

In particular, the bound $|e^{i\omega} - 1| \leq 2|\omega|^\alpha$ valid for all $\omega \in \mathbb{R}$ and $0 \leq \alpha \leq 1$ implies that

$$\begin{aligned} \left| \frac{\Phi(\kappa) - 1}{\kappa} \right| &\leq 2\pi\kappa^{\alpha-1} \iint_{\mathbb{R}^2} |x|^\alpha \phi(x+t) \psi(t) dx dt \\ &\leq 2\pi\kappa^{\alpha-1} \int |x|^\alpha \phi(x) dx \int |t|^\alpha \psi(t) dt \end{aligned}$$

Since both $\phi, \psi \in \mathcal{D}_\alpha$ for some $\alpha > 0$, this shows that the function $\frac{\Phi(\kappa) - \mathbb{1}_{\kappa \leq 1}}{\kappa}$ is integrable on \mathbb{R}_+ . Similarly, the function $\Psi(\kappa)/\kappa$ is also integrable on \mathbb{R}_+ . By the Riemann–Lebesgue lemma, this implies that for any $\omega \in \mathbb{R}$,

$$\int_0^\infty \sin(\epsilon\omega\kappa) \Psi(\kappa) \frac{d\kappa}{\kappa} = \mathcal{U}_\omega. \quad \epsilon \rightarrow 0 \quad (2.35)$$

By formula (2.31) and by definition of \hat{Q} , we have

$$\mathbb{E} [G_{\phi, \epsilon}(u) G_{\psi, \epsilon}(v)] = 2 \int_0^\infty \Re \left\{ e^{-2\pi i(u-v)\kappa} \hat{\phi}(\epsilon\kappa) \overline{\hat{\psi}(\epsilon\kappa)} \right\} \frac{\sin^2(\pi\ell\kappa)}{\kappa} d\kappa \quad (2.36)$$

$$\begin{aligned} &= \underbrace{\int_0^\infty 2 \cos\left(\frac{(u-v)\kappa}{\epsilon'}\right) \frac{\sin^2(\ell\kappa/2\epsilon')}{\kappa} \Phi(\kappa) d\kappa}_{:= I_1(u, v, \epsilon')} \\ &+ \underbrace{\int_0^\infty 2 \sin\left(\frac{(u-v)\kappa}{\epsilon'}\right) \frac{\sin^2(\ell\kappa/2\epsilon')}{\kappa} \Psi(\kappa) d\kappa}_{:= I_2(u, v, \epsilon')}. \end{aligned} \quad (2.37)$$

Using the trigonometric identity

$$-2 \cos(a) \sin^2(b/2) = 1 - \cos a - \frac{1 - \cos(b+a) + 1 - \cos(b-a)}{2},$$

with $a = (u-v)\kappa/\epsilon'$ and $b = \ell\kappa/\epsilon'$, we see that we can apply Lemma 2.10. We obtain for all $u, v \in \mathbb{R}$,

$$\begin{aligned} I_1(u, v, \epsilon') &= -\text{Cin}\left(\frac{u-v}{\epsilon'}\right) + \frac{1}{2} \left\{ \text{Cin}\left(\frac{\ell+u-v}{\epsilon'}\right) + \text{Cin}\left(\frac{\ell+v-u}{\epsilon'}\right) \right\} \\ &\quad - \mathcal{E}_\Phi\left(\frac{u-v}{\epsilon'}\right) + \frac{1}{2} \left\{ \mathcal{E}_\Phi\left(\frac{\ell+u-v}{\epsilon'}\right) + \mathcal{E}_\Phi\left(\frac{\ell+v-u}{\epsilon'}\right) \right\}, \end{aligned} \quad (2.38)$$

and the terms in (2.38) combine as the error term $\mathcal{U}_{u,v}$. Moreover, by formula (2.33), this implies that

$$I_1(u, v, \epsilon') = -\log^+\left(\frac{u-v}{\epsilon'}\right) + \frac{1}{2} \left\{ \log^+\left(\frac{\ell+u-v}{\epsilon'}\right) + \log^+\left(\frac{\ell+v-u}{\epsilon'}\right) \right\} + \mathcal{U}_{u,v},$$

By definition of the functions \log^+ and Q_ϵ , (2.32), this may be written as

$$I_1(u, v, \epsilon') = Q_\epsilon(u-v) + \mathcal{U}_{u,v}. \quad (2.39)$$

We can also evaluate the integral $I_2(u, v, \epsilon')$ using a similar argument. Since

$$2 \sin(a) \sin^2(b/2) = \sin a - \sin(a+b)/2 + \sin(a-b)/2,$$

the estimate (2.35) shows that $I_2(u, v, \epsilon') = \mathcal{U}_{u,v}$. Combining this fact with (2.39) and formula (2.37), this completes the proof. \square

Proposition 2.12. *Let $\phi, \psi \in \mathcal{D}$, we have for all $u, v \in \mathbb{R}$,*

$$\mathbb{E} [G_{\phi, \epsilon}(u) G_{\psi, \delta}(v)] = Q_{\epsilon \vee \delta}(u-v) + \mathcal{U}_{u,v},$$

where we can consider the limit as the parameters ϵ and δ converge to 0 in an arbitrary way.

Proof. We let $\epsilon' = \epsilon/2\pi$ and $\delta' = \delta/\epsilon$. Without loss of generality, we suppose that $\delta' \rightarrow \beta \in [0, 1]$ as $\epsilon \rightarrow 0$ (in particular, $\beta = 0$ if we consider successive limits as $\delta \rightarrow 0$ and $\epsilon \rightarrow 0$). By formula (2.36), we have

$$\mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \delta}(v)] = 2 \int_0^\infty \Re \left\{ e^{-(u-v)\kappa/\epsilon'} \hat{\phi}(\kappa) \overline{\hat{\psi}(\kappa\delta')} \right\} \frac{\sin^2(\ell\kappa/2\epsilon')}{\kappa} d\kappa.$$

In particular, this implies that

$$|\mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \delta}(v)] - \mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \epsilon\beta}(v)]| \leq 2 \int_0^\infty |\hat{\psi}(\kappa\beta) - \hat{\psi}(\kappa\delta')| |\hat{\phi}(\kappa)| \frac{d\kappa}{\kappa}.$$

Since $|\hat{\psi}(\kappa\beta) - \hat{\psi}(\kappa\delta')| \leq 4\pi|\kappa|^\alpha|\beta - \delta'|^\alpha \int |x|^\alpha \psi(x)dx$ for any $0 \leq \alpha \leq 1$, there exists a constant $C > 0$ which depends only on $\psi \in \mathcal{D}_\alpha$ so that if $0 < \alpha < 1/2$ is sufficiently small, then

$$|\mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \delta}(v)] - \mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \epsilon\beta}(v)]| \leq C|\beta - \delta'|^\alpha \int_0^\infty |\hat{\phi}(\kappa)| \frac{d\kappa}{\kappa^{1-\alpha}}. \quad (2.40)$$

Since $\phi \in L^1 \cap L^2(\mathbb{R})$, by the Cauchy–Schwarz inequality, the RHS of (2.40) is finite and converges to 0 as $\epsilon \rightarrow 0$ (by assumption, $\lim \delta' = \beta$). Thus, it suffices to prove that given $\beta \in [0, 1]$,

$$\mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \epsilon\beta}(v)] = Q_\epsilon(u - v) + \underset{\epsilon \rightarrow 0}{\mathcal{O}_{u,v}}. \quad (2.41)$$

If $\beta > 0$, this follows directly from Lemma 2.11, since $G_{\psi, \epsilon\beta} = G_{\psi\beta, \epsilon}$. If $\beta = 0$, since $\hat{\psi}(0) = 1$, we have

$$\mathbb{E}[G_{\phi, \epsilon}(u)G_{\psi, \epsilon\beta}(v)] = \int_0^\infty \Re \left\{ e^{-(u-v)\kappa/\epsilon'} \hat{\phi}(\kappa) \right\} \frac{\sin^2(\ell\kappa/2\epsilon')}{\kappa} d\kappa,$$

and the same computations as in the proof of Proposition 2.11 shows that the function $\Re \hat{\phi}$ satisfies the assumptions of Lemma 2.10 and the function $\kappa \mapsto \Im \hat{\phi}(\kappa)/\kappa$ is integrable on \mathbb{R}_+ , so that the asymptotics (2.40) hold. \square

Corollary 2.13. *Let $\ell > 0$, G be the stationary Gaussian process on \mathbb{R} with covariance function (1.22), and let $\phi \in \mathcal{D}$. For any $u, v \in \mathbb{R}$, let $T(u, v) = Q(u - v)$, and for any $\epsilon, \delta > 0$, define*

$$T_{\epsilon, \delta}(u, v) = \mathbb{E}[G_{\phi, \epsilon}(u)G_{\phi, \delta}(v)].$$

Then, the function $T_{\epsilon, \delta}$ satisfies Assumption 2.2.

Proof. By formula (2.32), $\lim_{\epsilon \rightarrow 0} Q_\epsilon(x) = Q(x)$ for almost all $x \in \mathbb{R}$, so that the condition (2.4) follows directly from Proposition 2.12. Similarly, (2.5) is also an immediate consequence Proposition 2.12. To get the upper-bound, we check that if $\epsilon \leq \ell \wedge 1$, directly from (2.32) we obtain

$$Q_\epsilon(x) \leq -\log(\epsilon \vee |x|) + \log(2\pi\sqrt{\ell^2 + |x|^2}).$$

We now use the identity $-\log^+(\epsilon^{-1} \wedge |x|^{-1}) = \log(\epsilon \vee |x|) + \log(|x|)\mathbb{1}_{|x| \geq 1}$, which follows from the definition (1.34) of \log^+ and the condition $\epsilon \leq 1$, to derive

$$\begin{aligned} Q_\epsilon(x) - \log^+(\epsilon^{-1} \wedge |x|^{-1}) \\ \leq \sup \left\{ \log(2\pi \sqrt{\ell^2 + |x|^2}) : |x| \leq 1 \right\} \\ \vee \sup \left\{ -\log |x| + \log(2\pi \sqrt{\ell^2 + |x|^2}) : |x| \geq 1 \right\}. \end{aligned}$$

Since the function $x \mapsto -\log |x| + \log(2\pi \sqrt{\ell^2 + |x|^2})$ is decreasing on the set $|x| \geq 1$, we have

$$Q_{\epsilon \vee \delta}(x) \leq \log^+(|x|^{-1} \wedge \epsilon^{-1} \wedge \delta^{-1}) + \log(2\pi \sqrt{\ell^2 + 1})$$

and the condition (2.3) also follows from Proposition 2.12. \square

3. Asymptotic Analysis

3.1. Proof of Proposition 1.10. The goal is to deduce the asymptotics of Proposition 1.10 from Theorem 1.13 by performing the asymptotics of the solution to the Riemann–Hilbert problem (1.44). Recall that for the sine process, we have

$$f(x) = \begin{pmatrix} e^{i\pi Nx} \\ e^{-i\pi Nx} \end{pmatrix}, \quad g(x) = \begin{pmatrix} e^{i\pi Nx} \\ -e^{-i\pi Nx} \end{pmatrix}, \quad (3.1)$$

so that we look for the asymptotics of the solution to the problem:

- $m(z)$ is analytic on $\mathbb{C} \setminus \mathbb{R}$.
- $m(z)$ satisfies the jump condition:

$$m_+(x) = m_-(x) \begin{pmatrix} 1 + \varphi(x) & -\varphi(x)e^{2\pi i Nx} \\ \varphi(x)e^{-2\pi i Nx} & 1 - \varphi(x) \end{pmatrix}, \quad x \in \mathbb{R} \quad (3.2)$$

- $m(z) \rightarrow I$ as $z \rightarrow \infty$ in $\mathbb{C} \setminus \mathbb{R}$.

Note that we do not emphasize that the matrix m depends on the dimension N to keep the notation simple. Moreover, the solution of (1.44) can be obtained from the solution of (3.2) simply by replacing φ by $\varphi_t = t\varphi$ for all $t \in [0, 1]$. Finally, we will generalize slightly the setting of Proposition 1.10 and work with the following assumptions.

Assumption 3.1. Suppose that φ is a function which is analytic in the strip $|\Im z| < \delta$, so that $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ and $\varphi \in L^1 \cap L^\infty(\mathbb{R} \pm is)$ for all $|s| \leq \delta/2$. In particular, we define

$$C_\varphi = \sup \left\{ \|\varphi\|_{L^\infty(\mathbb{R} + is)} \vee \|\varphi\|_{L^1(\mathbb{R} + is)} : s \leq \delta/2 \right\}.$$

Let us also assume that there exists a constant $c > 0$ so that

$$\inf_{|\Im z| < \delta, x < -1} |\varphi(z) - x| \geq c. \quad (3.3)$$

In particular, the function $\psi = \log(1 + \varphi)$ is also analytic in the strip $|\Im z| < \delta$. Finally, we assume that $\psi \in L^1 \cap L^\infty(\mathbb{R})$ and we let $C_\psi = c^{-1} \exp \|\psi\|_{L^\infty(\mathbb{R})}$.

Lemma 3.1. *Suppose that the function φ satisfies Assumption 3.1 and that $C_\varphi C_\psi^2 e^{-\pi\delta N} \rightarrow 0$ as $N \rightarrow \infty$. Then, the solution of the Riemann–Hilbert problem (3.2) satisfies for all $x \in \mathbb{R}$,*

$$m_+(x) = R(x) \begin{pmatrix} e^{\mathcal{C}(\psi)_+(x)} - \frac{\varphi(x)}{1+\varphi(x)} e^{\mathcal{C}(\psi)_+(x)+2\pi iNx} & \\ 0 & e^{-\mathcal{C}(\psi)_+(x)} \end{pmatrix}, \quad (3.4)$$

where $\mathcal{C}(\psi)$ denotes the Cauchy transform of the function $\psi = \log(1+\varphi)$ and the 2×2 matrix $R(z)$ is analytic in the strip $|\Im z| < \delta/4$ and satisfies the bound:

$$\|R(z) - I\| \ll \delta^{-1/2} C_\varphi C_\psi^2 e^{-\pi\delta N} \quad (3.5)$$

for the matrix supremum norm.

The proof of Lemma 3.1 will be given at the end of this section and it follows closely the proof of the Strong Szegő theorem given by Deift in [17, Example 3]. Given this result, let us first complete the proof of Proposition 1.10.

Proof of Proposition 1.10. First of all, we claim that, if the function $h(z)$ satisfies Assumption 1.2, then the function $\varphi_t(z) = t(e^{h(z)} - 1)$ satisfies Assumption 3.1 for all $t \in [0, 1]$. Indeed, we have

$$|\varphi_t(z)| \leq \left| \int_0^{h(z)} e^w dw \right| \leq C_\infty C_1$$

and for all $|s| \leq \delta/2$,

$$\int_{\mathbb{R}} |\varphi_t(x + is)| dx \leq C_\infty \int_{\mathbb{R}} |h(x + is)| dx \leq C_\infty C_1. \quad (3.6)$$

On the other hand, the condition $|\Im h(z)| < \alpha$ guarantees we can choose $c = |\sin \alpha|/2$ in (3.3). Thus, for all $t \in [0, 1]$, the functions $\psi_t(z) = \log(1 - t + te^{h(z)})$ are analytic in the strip $|\Im z| < \delta$, and since the log function is increasing on \mathbb{R}_+ , we have for all $x \in \mathbb{R}$,

$$\begin{aligned} \psi_t(x) \vee 0 &\leq \log(e^{h(x)} \vee 1) = h(x) \vee 0 \quad \text{and} \quad -\psi_t(x) \wedge 0 \leq -\log(e^{h(x)} \wedge 1) \\ &\leq h(x) \wedge 0. \end{aligned}$$

These inequalities show that

$$|\psi_t(x)| \leq |h(x)| \quad (3.7)$$

and it follows that $\psi_t \in L^1 \cap L^\infty(\mathbb{R})$ (in fact there is equality in (3.7) if and only if $t = 1$, in which case $\psi_1 = h$). The bottom line is that, for any $t \in [0, 1]$, the function $\varphi_t(z) = t(e^{h(z)} - 1)$ satisfies Assumption 3.1 with

$$C_{\varphi_t} = C_\infty C_1 \quad \text{and} \quad C_{\psi_t} = \frac{2C_\infty}{|\sin \alpha|}. \quad (3.8)$$

In the rest of the proof, we will denote for all $t \in [0, 1]$ and for all $x \in \mathbb{R}$,

$$\psi_t(z) = \log(1 + t\varphi) = \log(1 - t + te^{h(z)}) \quad \text{and} \quad H_t(x) = \mathcal{C}(\psi_t)_+(x).$$

By Theorem 1.13, formula (3.4) implies that

$$F_t(x) = m_+(x)f(x) = R_t(x)\tilde{F}_t(x) \quad \text{where} \quad \tilde{F}_t(x) = \begin{pmatrix} \frac{e^{i\pi Nx + H_t(x)}}{1+t\varphi(x)} \\ e^{-i\pi Nx - H_t(x)} \end{pmatrix},$$

and

$$G_t(x) = (m_+^{-1})^*(x)g(x) = (R_t(x)^{-1})^*\tilde{G}_t(x) \quad \text{where} \quad \tilde{G}_t(x) = \begin{pmatrix} e^{i\pi Nx + H_t(x)} \\ -\frac{e^{-i\pi Nx - H_t(x)}}{1+t\varphi(x)} \end{pmatrix}.$$

Since the function $h' \in L^1(\mathbb{R})$, the function H_t is continuously differentiable on \mathbb{R} for all $t \in [0, 1]$. Moreover, by the Plemelj-Sokhotski formula (A.6), we obtain

$$2H'_t = \frac{t\varphi'}{1+t\varphi} + i\mathcal{H}(\psi'_t). \quad (3.9)$$

So, if we differentiate the expression of $\tilde{F}_t(x)$, we obtain for all $x \in \mathbb{R}$,

$$\tilde{F}'_t(x) = \begin{pmatrix} (\pi i N + H'_t(x) - \frac{t\varphi'(x)}{1+t\varphi(x)}) \frac{1}{1+t\varphi(x)} e^{\pi i Nx + H_t(x)} \\ -(\pi i N + H'_t(x)) e^{-\pi i Nx - H_t(x)} \end{pmatrix}.$$

Since, $F_t^* G_t = 0$, this shows that

$$\begin{aligned} \left(\frac{d\sqrt{\varphi}F_t}{dx}(x) \right)^* \left(\sqrt{\varphi(x)}G_t(x) \right) &= \varphi(x)F'_t(x)^*G_t(x) \\ &= \frac{\varphi(x)}{1+t\varphi(x)} \left(-2\pi i N + 2\overline{H'_t(x)} - \frac{t\varphi'(x)}{1+t\varphi(x)} \right) + \varphi(x) (U_t(x)\tilde{F}_t(x))^* \tilde{G}_t(x), \end{aligned} \quad (3.10)$$

where $U_t(x) = R_t(x)^{-1}R'_t(x)$. Since the matrix $R_t(z)$ is analytic in the strip $|\Im z| < \delta/4$, by Cauchy's formula, $\|R'_t(x)\| \leq \delta^{-1}\|R_t(x) - I\|$. Thus, the estimate (3.5) and (3.8) show that

$$\|U_t(x)\| \ll \delta^{-3/2} C_{\varphi_t} C_{\psi_t} e^{-\pi\delta N} \ll \frac{C_\infty^3 C_1}{\delta^{3/2} |\sin \alpha|} e^{-\pi\delta N}.$$

On the other hand, $\Re\{H_t\} = \frac{1}{2} \log(1+t\varphi) = \frac{\psi_t}{2}$ so that, by the estimate (3.7), for all $x \in \mathbb{R}$ and all $t \in [0, 1]$,

$$\|\tilde{F}_t(x)\| \leq \sqrt{C_\infty} \quad \text{and} \quad \|\tilde{G}_t(x)\| \leq C_\infty^{3/2}.$$

This proves that the last term in formula (3.10) is bounded by

$$\left| \varphi(x) (U_t(x)\tilde{F}_t(x))^* \tilde{G}_t(x) \right| \ll |\varphi(x)| \frac{C_\infty^5 C_1}{\delta^{3/2} |\sin \alpha|} e^{-\pi\delta N}. \quad (3.11)$$

By (3.6), the RHS of (3.11) is integrable on $\mathbb{R} \times [0, 1]$ and its L^1 -norm contributes to the error term in formula (1.32). Therefore, by formula (1.43) and (3.10), in order to complete the proof, it remains to show that

$$\begin{aligned} &\int_0^1 \int_{\mathbb{R}} \frac{\varphi(x)}{1+t\varphi(x)} \left(2\pi i N - 2\overline{H'_t(x)} + \frac{t\varphi'(x)}{1+t\varphi(x)} \right) dx dt \\ &= 2\pi i N \int h(x) dx + i\pi \|h\|_{H^{1/2}(\mathbb{R})}^2. \end{aligned} \quad (3.12)$$

This is a rather easy computation that we aligned in two steps. First observe that, since $\varphi(x) = e^{h(x)} - 1$, we have

$$\int_0^1 \frac{\varphi(x)}{1+t\varphi(x)} dt = [\log(1+t\varphi(x))]_{t=0}^1 = h(x).$$

By Fubini's theorem, this yields the first term in formula (3.12). Secondly, by formula (3.9), we have

$$\frac{\varphi}{1+t\varphi} \left(2H'_t - \frac{t\varphi'}{1+t\varphi} \right) = i \frac{d\psi_t}{dt} \mathcal{H} \left(\frac{d\psi_t}{dx} \right).$$

By Plancherel's formula and the linearity of the Fourier transform, this implies that

$$\begin{aligned} \int_{\mathbb{R}} \frac{d\psi_t}{dt}(x) \mathcal{H} \left(\frac{d\psi_t}{dx} \right)(x) dx &= 2\pi \int_{\mathbb{R}} |\kappa| \frac{d\widehat{\psi}_t}{dt}(\kappa) \overline{\widehat{\psi}_t(\kappa)} d\kappa \\ &= \pi \frac{d}{dt} \left(\underbrace{\int_{\mathbb{R}} |\kappa| \widehat{\psi}_t(\kappa) \overline{\widehat{\psi}_t(\kappa)} d\kappa}_{= \|\psi_t\|_{H^{1/2}(\mathbb{R})}^2} \right) \end{aligned}$$

Note that we can pull the differential $\frac{d}{dt}$ out of the integral because both functions $\psi_t, \frac{d\psi_t}{dt} \in L^2(\mathbb{R})$ by our assumptions. We conclude that

$$\int_0^1 \int_{\mathbb{R}} \frac{\varphi(x)}{1+t\varphi(x)} \left(2H'_t(x) - \frac{t\varphi'(x)}{1+t\varphi(x)} \right) dx dt = i\pi \left(\|\psi_1\|_{H^{1/2}(\mathbb{R})}^2 - \|\psi_0\|_{H^{1/2}(\mathbb{R})}^2 \right).$$

Since $\psi_1 = h$ and $\psi_0 = 0$, this yields the second term in formula (3.12). \square

In order to get the uniform estimate (3.5) in Lemma 3.1, we will need the following correspondence between a well-posed RHP and a singular integral equation.

Theorem 3.2 (Kuijlaars [45], Theorem 3.1). *Let Σ be an oriented contour in the complex plane and let $\Delta \in L^2 \cap L^\infty(\Sigma)$ be a $n \times n$ matrix. Suppose that R is a $n \times n$ matrix which is analytic in $\mathbb{C} \setminus \Sigma$ and satisfies*

$$\begin{cases} R_+ = R_-(I + \Delta) & \text{on } \Sigma \\ R(z) \rightarrow I & \text{as } z \rightarrow \infty \end{cases}. \quad (3.13)$$

We associate to Δ an operator \mathcal{C}_Δ acting on $n \times n$ matrices in $L^2(\Sigma)$ and defined by

$$\mathcal{C}_\Delta Y(z) = \lim_{w \rightarrow z_+} \frac{1}{2\pi i} \int_{\Sigma} \frac{Y(s)\Delta(s)}{s-w} ds.$$

Suppose that $\|\Delta\|_{L^\infty(\Sigma)}$ is sufficiently small so that the equation

$$X - \mathcal{C}_\Delta X = \mathcal{C}_\Delta I \quad (3.14)$$

has a unique solution $X \in L^2(\Sigma)$. Then, there exists a constant $C > 0$ which only depends on the contour Σ such that

$$\|X\|_{L^2(\Sigma)} \leq \frac{C\|\Delta\|_{L^2(\Sigma)}}{1 - C\|\Delta\|_{L^\infty(\Sigma)}}, \quad (3.15)$$

and the RHP (3.13) has a unique solution which is given by

$$R(z) = I + \frac{1}{2\pi i} \int_{\Sigma} \frac{(I + X(s))\Delta(s)}{s - z} ds$$

for any $z \in \mathbb{C} \setminus \Sigma$.

Armed with the above theorem, we can construct the solution of the Riemann–Hilbert problem (3.2) and check the estimate (3.5).

Proof of Lemma 3.1. If $\psi(x) = \log(1 + \varphi(x))$, then we can rewrite the jump matrix as

$$v(x) = \begin{pmatrix} e^{\psi(x)} & -\varphi(x)e^{2\pi i N x} \\ \varphi(x)e^{-2\pi i N x} & (1 - \varphi(x)^2)e^{-\psi(x)} \end{pmatrix}.$$

In order to apply the Deift–Zhou steepest method, we use the decomposition:

$$v(x) = \underbrace{\begin{pmatrix} 1 & 0 \\ \varphi(x)e^{-\psi(x)-2\pi i N x} & 1 \end{pmatrix}}_{= \tilde{A}(x)} \underbrace{\begin{pmatrix} e^{\psi(x)} & 0 \\ 0 & e^{-\psi(x)} \end{pmatrix}}_{= \Psi(x)} \underbrace{\begin{pmatrix} 1 - \varphi(x)e^{-\psi(x)+2\pi i N x} \\ 0 & 1 \end{pmatrix}}_{= A(x)}. \quad (3.16)$$

By assumption, the matrices A and \tilde{A} are analytic in the domain $|\Im z| < \delta$ and we can define a matrix M on $\mathbb{C} \setminus \{\mathbb{R} \cup \Gamma_{\pm}\}$ by setting:

$$\begin{array}{c} \textcolor{red}{M(z) = m(z)} \\ \hline \textcolor{red}{\Gamma_+ = \mathbb{R} + i\delta/2} \\ M(z) = m(z)A^{-1}(z) \\ \hline 0 \\ \hline M(z) = m(z)\tilde{A}(z) \\ \hline \textcolor{red}{\Gamma_- = \mathbb{R} - i\delta/2} \\ \textcolor{red}{M(z) = m(z)} \end{array}$$

We deduce from the Riemann–Hilbert problem (3.2), that the matrix M has the following properties:

- $M(z)$ is analytic on $\mathbb{C} \setminus (\Gamma_{\pm} \cup \mathbb{R})$.
- $M(z)$ satisfies the following jump conditions

$$\begin{cases} M_+(z) = M_-(z)A(z), & z \in \Gamma_+ \\ M_+(x) = M_-(x)\Psi(x), & x \in \mathbb{R} \\ M_+(z) = M_-(z)\tilde{A}(z), & z \in \Gamma_- \end{cases} \quad (3.17)$$

- $M(z) \rightarrow I$ as $z \rightarrow \infty$.

Since $\psi \in L^1(\mathbb{R})$, its Cauchy transform is well-defined, c.f. (A.5), and we claim that the global parametrix is given by

$$P(z) = \begin{pmatrix} e^{\mathcal{C}(\psi)(z)} & 0 \\ 0 & e^{-\mathcal{C}(\psi)(z)} \end{pmatrix}. \quad (3.18)$$

Indeed, it is straightforward to check using formula (A.6), that it solves the RHP:

- $P(z)$ is analytic on $\mathbb{C} \setminus \mathbb{R}$.
- $P(z)$ satisfies the condition $P_+ = P_- \Psi$ on \mathbb{R}
- $P(z) \rightarrow I$ as $z \rightarrow \infty$.

The matrix $P(z)$ is invertible on $\mathbb{C} \setminus \mathbb{R}$ and, if we let $R = MP^{-1}$, then the matrix R solves the RHP:

- $R(z)$ is analytic on $\mathbb{C} \setminus (\mathbb{R} \cup \Gamma_{\pm})$.
- If we let $\Delta = PAP^{-1} - I$ and $\tilde{\Delta} = P\tilde{A}P^{-1} - I$, then $R(z)$ satisfies the jump conditions

$$\begin{cases} R_+ = R_- (I + \Delta) & \text{on } \Gamma_+ \\ R_+ = R_- (I + \tilde{\Delta}) & \text{on } \Gamma_- \end{cases} \quad (3.19)$$

- $R(z) \rightarrow I$ as $z \rightarrow \infty$.

Moreover observe that by formulae (3.16) and (3.18),

$$\begin{aligned} \Delta(z) &= \begin{pmatrix} 0 & -\varphi(z)e^{-\psi(z)+2\pi i N z + 2\mathcal{C}(\psi)(z)} \\ 0 & 0 \end{pmatrix}, & \forall z \in \Gamma_+ \\ \tilde{\Delta}(z) &= \begin{pmatrix} 0 & 0 \\ \varphi(z)e^{-\psi(z)-2\pi i N z - 2\mathcal{C}(\psi)(z)} & 0 \end{pmatrix}, & \forall z \in \Gamma_-. \end{aligned}$$

The function ψ is real valued on \mathbb{R} and for any $z \in \Gamma_{\pm}$, we have

$$\Re\{\mathcal{C}(\psi)(z)\} = \frac{\delta}{4\pi} \int_{\mathbb{R}} \frac{\psi(x)}{(\Re z - x)^2 + \delta^2/4} dx,$$

so that

$$|\Re\{\mathcal{C}(\psi)(z)\}| \leq \|\psi\|_{L^\infty(\mathbb{R})}/2.$$

Combined with Assumption 3.1, this implies that for any $z \in \Gamma_+$,

$$\|\Delta(z)\| = \left| \frac{\varphi(z)}{1 + \varphi(z)} \right| e^{\Re\{2\pi i N z + 2\mathcal{C}(\psi)(z)\}} \leq |\varphi(z)| C_\psi^2 e^{-\pi\delta N},$$

so that the matrix $\Delta \in L^1 \cap L^\infty(\Gamma_+)$ and

$$\|\Delta\|_{L^\infty(\Gamma_+)} \vee \|\Delta\|_{L^1(\Gamma_+)} \leq C_1 C_\psi^2 e^{-\pi\delta N}. \quad (3.20)$$

Note that, since $\tilde{\Delta}(z) = -\Delta(\bar{z})^*$ for any $z \in \Gamma_-$, the matrix $\tilde{\Delta}$ also satisfies the estimate (3.20). Hence, by Theorem 3.2, we obtain that the solution of the RHP (3.19) is unique and is given by

$$R(z) = I + \frac{1}{2\pi i} \int_{\Gamma_+} \frac{(I + X(s))\Delta(s)}{s - z} ds + \frac{1}{2\pi i} \int_{\Gamma_-} \frac{(I + \tilde{X}(s))\tilde{\Delta}(s)}{s - z} ds \quad (3.21)$$

where the 2×2 matrices X and \tilde{X} solve appropriate singular integral equations, c.f. (3.14). Moreover, a simple estimate using the Cauchy–Schwarz inequality shows that for any $z \in \mathbb{C}$ such that $|\Im z| \leq \delta/4$,

$$\left\| \frac{1}{2\pi i} \int_{\Gamma_+} \frac{(I + X(s))\Delta(s)}{s - z} ds \right\| \leq \frac{1}{2\sqrt{\pi}\delta} (\|\Delta\|_{L^2(\Gamma_+)} + \|\Delta\|_{L^\infty(\Gamma_+)} \|X\|_{L^2(\Gamma_+)}).$$

Hence, as $\|\Delta\|_{L^\infty(\Gamma_+)} \rightarrow 0$ as $N \rightarrow \infty$, combining the bounds (3.15) and (3.20), we have proved that

$$\left\| \frac{1}{2\pi i} \int_{\Gamma_+} \frac{(\mathbf{I} + X(s))\Delta(s)}{s - z} ds \right\| \ll \delta^{-1/2} C_\psi^2 C_\varphi e^{-\pi\delta N}.$$

A similar estimate holds for the integral over Γ_- and, by formula (3.21), this proves the bound (3.5). Going back to the original problem, we see that for all $x \in \mathbb{R}$, $m_+(x) = M_+(x)A(x)$ where $M_+(x) = R(x)P_+(x)$. Finally, we deduce (3.4) from formulae (3.16) and (3.18). \square

3.2. Strong Gaussian approximation for the CUE. The goal of this section is to prove the Gaussian asymptotics (1.20) for the CUE. We work with the statistic (1.7) and it will be convenient to use the notation

$$h_{\mathbf{u}, \epsilon}(x) = \gamma\pi \sum_{k=1}^q t_k \int_{u_k - \ell/2}^{u_k + \ell/2} \phi_{\epsilon_k}(x - t) dt \quad (3.22)$$

We shall suppose that the mollifier ϕ belongs to the Schwartz class, i.e. ϕ is a smooth function such that for all $\eta > 0$ we have $\phi(x) = O(|x|^{-\eta})$ as $x \rightarrow \pm\infty$. The main goal of this section will be to prove

Proposition 3.3. *Consider a CUE matrix U of size $N \times N$ with eigenangles $\theta_1, \dots, \theta_N \in [-\pi, \pi)$ and mesoscopic scale $0 < \alpha < 1$. Take $\epsilon^* = \min\{\epsilon_1, \dots, \epsilon_q\}$. Let $\delta > 0$ and suppose that we have the following bound $N^\alpha/(N\epsilon^*) = O(N^{-\delta})$. Then for all Schwartz ϕ , the following Gaussian estimate holds as $N \rightarrow \infty$*

$$\begin{aligned} & \mathbb{E} \left[\exp \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}(N^\alpha \theta_j) - \mathbb{E} \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}(N^\alpha \theta_j) \right) \right) \right] \\ &= \exp \left(\frac{\|h_{\mathbf{u}, \epsilon}\|_{H^{1/2}}^2}{2} \right) (1 + \mathcal{E}_N^S) e^{\mathcal{E}_N^{\text{glob}}} \end{aligned} \quad (3.23)$$

where for any $\eta > 0$, the smoothing error term satisfies the bound

$$|\mathcal{E}_N^S| \leq C\gamma^2 \|t\|_1^{\eta_2} N^{-\eta} \exp \left(\gamma^2 \|t\|_1^{\eta_2} N^{-\eta} \right) \quad (3.24)$$

for some $\eta_2 > 0$, while the global error term satisfies

$$|\mathcal{E}_N^{\text{glob}}| \leq C\|t\|_1^2 \log(\ell/\epsilon^*) \ell N^{-\alpha} \quad (3.25)$$

uniformly in compact subsets of the parameters $u_1, \dots, u_q \in \mathbb{R}$.

With the Gaussian approximation above we may complete the proofs of the main theorems for the CUE.

Proof of Theorems 1.3 and 1.6. Proposition 3.3 shows that the regularized statistic (1.7) satisfies the strong Gaussian approximation as in Assumption 2.4 for all $q \in \mathbb{N}$ (hence Assumption 2.5 holds as well). The fact that the associated covariance kernel satisfies Assumption 2.2 is a consequence of the computations carried out in Sect. 2.3, namely Theorem 2.11. Therefore we deduce Theorems 1.3 and 1.6 as corollaries of Theorems 2.6 and 2.1 respectively. \square

Remark 3.4. The error (3.24) is called the smoothing error because it fails when $\epsilon \rightarrow 0$ quickly enough that $\epsilon \sim N^\alpha/N$. Beyond this regime the asymptotics in (3.23) are no longer valid and one enters the transition regime of Fisher–Hartwig asymptotics, see [18, 44] for a review. On the other hand, (3.25) is controlled by the relative numbers of eigenvalues sampled by the statistic and is only small in the mesoscopic regime $0 < \alpha < 1$.

As for the sine process and Proposition 1.10, the proof of Proposition 3.3 is also based on the existence of integrable and determinantal structures in the CUE. In fact the Laplace transform of (1.7) could also be written as a Fredholm determinant which could then be analysed with an appropriate Riemann–Hilbert problem (as was done in [17] for the global scale $\alpha = 0$ and $\epsilon > 0$ fixed). However, for the CUE we will give a more elementary proof that does not involve Riemann–Hilbert techniques, but instead relies on an ‘algebraic miracle’ known (for historical reasons) as the Borodin–Okounkov–Case–Geronomo formula. Unlike our Riemann–Hilbert computation, the proof we give here does not impose any analyticity condition on the mollifier ϕ , but we do require it to be smooth in general. Another crucial difference is that the CUE has a macroscopic regime (defined by (1.7) with $\alpha = 0$) which has no analogue for the sine process.

When working with the CUE it is convenient to expand such functions in a Fourier series. Therefore in what follows we are going to work with the periodisation

$$h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta) = \sum_{a=-\infty}^{\infty} h_{\mathbf{u}, \epsilon}(N^\alpha(\theta + 2\pi a)). \quad (3.26)$$

Then $h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta + 2\pi) = h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta)$ is 2π -periodic. The periodisation has the convenient property that its Fourier coefficients are given explicitly in terms of the Fourier transform of $h_{\mathbf{u}, \epsilon}$.

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta) e^{-ik\theta} d\theta = N^{-\alpha} \int_{\mathbb{R}} h_{\mathbf{u}, \epsilon}(2\pi x) e^{-2i\pi kx N^{-\alpha}} dx \quad (3.27)$$

Furthermore, the linear statistics of $h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta)$ are uniformly close to those of $h_{\mathbf{u}, \epsilon}(\theta)$. This follows from the rapid decay of ϕ , since the difference between the two is deterministically bounded by

$$\sum_{|a|>0} \sum_{j=1}^N \sum_{k=1}^q \epsilon_k^{\eta-1} \int_{x_j - u_k - l/2}^{x_j - u_k + l/2} |s|^{-\eta} ds \leq \|t\| c N(\epsilon_*/N^\alpha)^\eta \quad (3.28)$$

where $x_j = N^\alpha(\theta_j + 2\pi a)$ and we used that $lN^{-\alpha} \rightarrow 0$ as $N \rightarrow \infty$. Choosing $\eta > 0$ large enough we can always ensure that (3.28) goes to 0 as $N \rightarrow \infty$. Hence it will suffice to always work with the periodisation.

Proposition 3.5 (Macroscopic approximation). *Define the quantity*

$$E_{\mathbf{u}, \epsilon} := \frac{1}{(2\pi)^2} N^{-\alpha} \sum_{k=1}^{\infty} k N^{-\alpha} |\hat{h}_{\mathbf{u}, \epsilon}(k/(2\pi N^\alpha))|^2 \quad (3.29)$$

and suppose that the hypotheses of Proposition 3.3 are satisfied. Then we have the Gaussian estimate

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j) - \mathbb{E} \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j) \right) \right) \right] = \exp(E_{\mathbf{u}, \epsilon}) (1 + \mathcal{E}_N^S) \quad (3.30)$$

where the error term \mathcal{E}_N^S satisfies (3.24) and is uniform in the variables u_1, \dots, u_q varying in compact subsets of \mathbb{R} .

To prove Proposition 3.5, the main idea is to exploit the fact that for the CUE, the left-hand side of (3.30) can be written exactly as an $N \times N$ Toeplitz determinant involving the Fourier coefficients of the periodic function $w_{\mathbf{u}, \epsilon}(\theta) = e^{h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta)}$. The representation as a determinant follows from the following well known but remarkable chain of equalities for the (un-centered) left-hand side of (3.23):

$$\mathbb{E} \left[\exp \left(\sum_{j=1}^q X_{N, \epsilon_N}(u_j) \right) \right] \quad (3.31)$$

$$= \frac{1}{(2\pi)^N N!} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \prod_{j=1}^N w(\theta_j) \prod_{1 \leq p < q \leq N} |e^{i\theta_q} - e^{i\theta_p}|^2 d\theta_1 \dots d\theta_N \quad (3.32)$$

$$= \frac{1}{(2\pi)^N N!} \int_{-\pi}^{\pi} \dots \int_{-\pi}^{\pi} \det\{w(\theta_j) e^{i(k-1)\theta_j}\}_{j,k=1}^N \det\{e^{-i(k-1)\theta_j}\}_{j,k=1}^N d\theta_1 \dots d\theta_N \quad (3.33)$$

$$= \det \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} w(\theta) e^{-i(k-j-2)\theta} d\theta \right\}_{j,k=1}^N \quad (3.34)$$

$$= \det T_N(w) \quad (3.35)$$

where $T_N(w)$ is the $N \times N$ Toeplitz matrix $\{\hat{w}_{k-j-2}\}_{j,k=1}^N$. That (3.31) equals (3.32) is a consequence of the Weyl integration formula for the unitary group. Then (3.33) writes the product of differences in (3.32) as a product of two determinants and finally (3.34) is a consequence of the Andrejeff identity.

Thus our task will be to calculate the asymptotics of the Toeplitz determinant in (3.35) as $N \rightarrow \infty$. The function $w(\theta)$ is called the *symbol*. For smooth and N -independent symbols, such asymptotics are well known from the strong Szegő limit theorem. However, our symbol is N -dependent and furthermore the quantity $E_{\mathbf{u}, \epsilon}$ is divergent in the limit $N \rightarrow \infty$. This is due to the fact that our symbol becomes discontinuous as $\epsilon_N \rightarrow 0$ and therefore does not belong to the $H^{1/2}$ -space. The following formula is our main tool for establishing the strong Gaussian approximation in the CUE.

Theorem 3.6 (Borodin–Okounkov–Case–Geronimo formula). *Let $w(\theta)$ a periodic function on the interval $\theta \in [0, 2\pi)$ and that $\log w(\theta)$ has Fourier coefficients \hat{L}_k . Suppose that we have the expansion*

$$\log w(\theta) = \sum_{k=-\infty}^{\infty} \hat{L}_k e^{ik\theta}, \quad \sum_{k=1}^{\infty} k |\hat{L}_k|^2 < \infty \quad (3.36)$$

In terms of the quantities $b(\theta) = 1/c(\theta)$ and

$$c(\theta) = \exp \left(\sum_{k=1}^{\infty} \hat{L}_k e^{ik\theta} - \sum_{k=1}^{\infty} \hat{L}_{-k} e^{-ik\theta} \right) \quad (3.37)$$

we have the following identity

$$\frac{\det(T_N(w))}{\exp \left(N \hat{L}_0 + \sum_{k=1}^{\infty} k |\hat{L}_k|^2 \right)} = \det(I - R_N H(b) H(\tilde{c}) R_N) \quad (3.38)$$

where R_N is the projection operator on $\ell^2(N+1, N+2, \dots)$ and $H(b)$, $H(\tilde{c})$ are infinite Hankel matrices corresponding to the sequence of Fourier coefficients of $b(\theta)$ and $c(\theta)$,

$$H(b) = \begin{pmatrix} b_1 & b_2 & b_3 & \dots \\ b_2 & b_3 & b_4 & \dots \\ b_3 & b_4 & b_5 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}, \quad H(\tilde{c}) = \begin{pmatrix} c_{-1} & c_{-2} & c_{-3} & \dots \\ c_{-2} & c_{-3} & c_{-4} & \dots \\ c_{-3} & c_{-4} & c_{-5} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \quad (3.39)$$

Proof. There are many proofs in the literature, for example the one of Borodin and Okounkov [9] led to an increased interest in the formula. For a comprehensive proof under the conditions mentioned here see Simon [66, Theorem 6.2.14], which also provides a detailed historical discussion of (3.38) and its original discovery in [33]. \square

Note that here $\log w(\theta) = h_{\mathbf{u}, \epsilon}^{(2\pi)}$ and the quantity $\sum_{k=1}^{\infty} k |\hat{L}_k|^2$ in (3.38) is precisely $E_{\mathbf{u}, \epsilon}$ in (3.29). Furthermore, an easy computation shows that $N \hat{L}_0 = \mathbb{E}(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j))$ which corresponds to a re-centering by the expectation. Hence formula (3.38) implies

$$\begin{aligned} & \mathbb{E} \left[\exp \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j) - \mathbb{E} \left(\sum_{j=1}^N h_{\mathbf{u}, \epsilon}^{(2\pi)}(\theta_j) \right) \right) \right] \\ &= \exp(E_{\mathbf{u}, \epsilon}) \det(I - R_N H(b) H(\tilde{c}) R_N). \end{aligned} \quad (3.40)$$

The next lemma gives us the necessary estimate on the above determinant, thus concluding the proof of Proposition 3.5.

Lemma 3.7. *Suppose the hypotheses of Proposition 3.3 are satisfied. Then for any $\eta > 0$, there exists $\eta_2 > 0$ such that we have the following estimate*

$$|\det(1 - R_N H(b) H(\tilde{c}) R_N) - 1| \leq C \gamma^2 \|t\|_1^{\eta_2} N^{-\eta} \exp \left(\gamma^2 \|t\|_1^{\eta_2} N^{-\eta} \right) \quad (3.41)$$

uniformly for variables u_1, \dots, u_q belonging to a compact subset of \mathbb{R} .

Proof. The following inequality is an easy consequence of standard properties of the Fredholm determinant

$$|\det(I + A) - 1| \leq e^{\|A\|_1} - 1 \quad (3.42)$$

In our case $A = -R_N H(b) H(\tilde{c}) R_N = -|H(\tilde{c}) R_N|^2 \leq 0$ using the fact that $H(b) = H(\tilde{c})^\dagger$, so $-A$ is a positive operator. Thus we can write the bound (3.42) in terms of the Hilbert–Schmidt norm

$$|\det(I - R_N H(b) H(\tilde{c}) R_N) - 1| \leq e^{\|H(\tilde{c}) R_N\|_2^2} - 1 \leq \|H(\tilde{c}) R_N\|_2^2 e^{\|H(\tilde{c}) R_N\|_2^2} \quad (3.43)$$

which can be computed explicitly (see 6.2.57 in [66])

$$\|H(\tilde{c})R_N\|_2^2 = \sum_{k=1}^{\infty} k|c_{k+N}|^2. \quad (3.44)$$

We now proceed to estimate the Fourier coefficients of the function $c(\theta)$ in (3.37), which clearly satisfies $|c(\theta)| = 1$. We have

$$c_{k+N} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(k+N)\theta} e^{2iu(\theta)} d\theta \quad (3.45)$$

where

$$u(\theta) = \Im \left\{ \sum_{k=1}^{\infty} \hat{L}_k e^{ik\theta} - \sum_{k=1}^{\infty} \hat{L}_{-k} e^{-ik\theta} \right\}. \quad (3.46)$$

The idea is to exploit cancellations in the integral (3.45) for large N coming from rapid oscillations of the factor $e^{-i(k+N)\theta}$. To this end, we integrate by parts p times, obtaining

$$c_{k+N} = \frac{2^p}{2\pi} \frac{1}{(k+N)^p} \int_{-\pi}^{\pi} e^{-i(k+N)\theta} e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right) e^{2iu(\theta)} d\theta \quad (3.47)$$

The function $e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right)$ is a polynomial in $u(\theta)$ and all its derivatives up to order p . We have the explicit formula

$$e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right) = \sum_{m=1}^p \sum_{\substack{k_1+2k_2+\dots+pk_p=p \\ k_1+k_2+\dots+k_p=m}} c_{p,\mathbf{k}} \prod_{l=1}^p u^{(l)}(\theta)^{k_l} \quad (3.48)$$

where $c_{p,\mathbf{k}}$ are some combinatorial coefficients. We now proceed to estimate $u^{(l)}(\theta)$. Clearly the interchange of derivative and sum in (3.46) is valid as the partial sums of all derivatives are uniformly convergent. To calculate the coefficients c_{k+N} , note that by (3.27) and the convolution theorem, we have

$$\begin{aligned} |u^{(l)}(\theta)| &\leq 2 \sum_{k=1}^{\infty} k^l |\hat{L}_k| \\ &\leq \frac{2\pi\gamma}{N^\alpha} \sum_{k=1}^{\infty} k^l \sum_{j=1}^q |t_j| \left| \frac{e^{-2i\pi(u_j+l/2)kN^{-\alpha}} - e^{-2i\pi(u_j-l/2)kN^{-\alpha}}}{2\pi k N^{-\alpha}} \right| |\hat{\phi}(\epsilon_j k / (2\pi N^\alpha))| \end{aligned} \quad (3.49)$$

$$\leq 2\gamma \sum_{j=1}^q \rho_j^{l-1} \sum_{k=1}^{\infty} (k/\rho_j)^{l-1} |t_j| |\hat{\phi}(k/(2\pi\rho_j))| \quad (3.50)$$

$$\sim 2\gamma \sum_{j=1}^q |t_j| \rho_j^l \int_0^{\infty} k^{l-1} |\hat{\phi}(k/(2\pi))| dk \quad (3.51)$$

where $\rho_j = N^\alpha/\epsilon_j \rightarrow \infty$. Hence there exists a constant $C > 0$ such that

$$|u^{(l)}(\theta)| \leq C \sum_{j=1}^q \rho_j^l |t_j| \leq C(N^\alpha/\epsilon^*)^l \sum_{j=1}^q |t_j| \quad (3.52)$$

uniformly in u_1, \dots, u_q . Inserting (3.52) into (3.48) yields the estimate

$$\left| e^{-u(\theta)} \left(\frac{d^p}{d\theta^p} e^{u(\theta)} \right) \right| \leq \sum_{m=1}^p \sum_{\substack{k_1+2k_2+\dots+pk_p=p \\ k_1+k_2+\dots+k_p=m}} c_{p,k} \prod_{l=1}^p (\|t\| N^\alpha / \epsilon^*)^{lk_l} = C(\|t\| N^\alpha / \epsilon^*)^p \quad (3.53)$$

Inserting this into (3.47) gives the following bound on the Fourier coefficients of $c(\theta)$

$$|c_{k+N}| \leq C\gamma \left(\frac{\|t\| N^\alpha / \epsilon^*}{k+N} \right)^p \quad (3.54)$$

and the corresponding bound on the Hilbert–Schmidt norm

$$\begin{aligned} \sum_{k=1}^{\infty} k |c_{k+N}|^2 &\leq C\gamma \sum_{k=1}^{\infty} k \left(\frac{\|t\| N^\alpha / \epsilon^*}{k+N} \right)^{2p} \leq C \int_1^{\infty} k \left(\frac{\|t\| N^\alpha / \epsilon^*}{k+N} \right)^{2p} dk \\ &= O(N^2 (\|t\| N^\alpha / (\epsilon^* N))^{2p}). \end{aligned} \quad (3.55)$$

This completes the proof of the lemma. \square

The quantity (3.29) is close to a Riemann sum. We need good estimates on the error in the approximation, which is the purpose of the next lemma. This error is generically of order $N^{-\alpha}$ in the mesoscopic regime. More generally, for the case of diverging interval length $l = L(N) \rightarrow \infty$ the error becomes of order $\log(L(N))L(N)N^{-\alpha}$.

Proposition 3.8 (Macroscopic to mesoscopic). *Consider the quantity (3.29). We have the uniform approximation*

$$E_{\mathbf{u}, \epsilon} = \frac{1}{(2\pi)^2} \int_0^\infty k |\hat{h}_{\mathbf{u}, \epsilon}(k)|^2 dk + O(\log(L(N))L(N)N^{-\alpha}) \quad (3.56)$$

Proof. We use the fact that the error in a Riemann sum approximation is given by the step size (here $N^{-\alpha}$) multiplied by the total variation norm of the function in question. Hence we have to estimate the quantity

$$\mathcal{E} := N^{-\alpha} \int_0^\infty \left| \frac{d}{dk} \left(k \left| \sum_{j=1}^q t_j \frac{e^{-2\pi i k(u_j + l/2)} - e^{-2\pi i k(u_j - l/2)}}{2\pi k} \hat{\phi}(k\epsilon_j) \right|^2 \right) \right| dk \quad (3.57)$$

Using the identity

$$\begin{aligned} &\left| \sum_{j=1}^q \hat{\phi}(k\epsilon_j) t_j (e^{-2\pi i k(u_j + l/2)} - e^{-2\pi i k(u_j - l/2)}) \right|^2 \\ &= 8 \sin^2(2\pi kl/2) \sum_{j_1 \leq j_2} t_{j_1} t_{j_2} \hat{\phi}(k\epsilon_{j_1}) \hat{\phi}(-k\epsilon_{j_2}) \cos(2\pi k(u_{j_1} - u_{j_2})) \end{aligned} \quad (3.58)$$

and changing variables $k \rightarrow k/\epsilon^*$, we see that it is sufficient to bound the quantity

$$I_{\epsilon, u} := \int_0^\infty \left| \frac{d}{dk} \frac{\sin^2(\pi kl) \cos(2\pi k(u_{j_2} - u_{j_1}))}{k} \hat{\phi}(k\epsilon_{j_1}) \hat{\phi}(-k\epsilon_{j_2}) \right| dk \quad (3.59)$$

uniformly in u as $\epsilon \rightarrow 0$. Computing the derivative yields four terms $I = I_1 + I_2 + I_3 + I_4$ coming from differentiating $1/k$, the two trigonometric terms and the functions $\hat{\phi}$, respectively. The contribution coming from the derivative of $1/k$ is bounded by

$$I_1 \leq \int_0^\infty \frac{\sin^2(\pi kl)}{k^2} dk = \frac{\pi^2}{2} l \quad (3.60)$$

To compute the contribution coming from the trigonometric terms we change variables $k \rightarrow k/\epsilon^*$ so that the argument of the Fourier transform is diverging for $k > 1$. Then the contribution is dominated by the interval $k \in [0, 1]$ due to the rapid decay of $\hat{\phi}(k)$ and we get the bounds

$$\begin{aligned} I_2 &\leq Cl \int_0^1 \left| \frac{\sin(2\pi kl/\epsilon^*)}{k} \right| dk = O(l \log(l/\epsilon^*)) \\ I_3 &\leq C2\pi |u_{j_2} - u_{j_1}| \int_0^1 \frac{\sin^2(\pi kl/\epsilon^*)}{k} dk = O(|u_{j_2} - u_{j_1}| \log(l/\epsilon^*)) \end{aligned} \quad (3.61)$$

A similar estimate yields

$$I_4 \leq \max\{\epsilon_{j_1}, \epsilon_{j_2}\} \int_0^1 \frac{\sin^2(\pi kl/\epsilon^*)}{k} dk = O(\max\{\epsilon_{j_1}, \epsilon_{j_2}\} \log(l/\epsilon^*)) \quad (3.62)$$

Multiplying these estimates by the step size $N^{-\alpha}$ we get the error in the Riemann sum approximation

$$\mathcal{E} \leq CN^{-\alpha} \sum_{j_1 \leq j_2} |t_{j_2} t_{j_1}| l \log(l/\epsilon^*) \quad (3.63)$$

which completes the proof of the Proposition. \square

A. Properties of the Hilbert Transform, $H^{1/2}$ -Noise, and a Log-Correlated Gaussian Process

We define the Fourier transform of any function $f \in L^1(\mathbb{R})$ by

$$\mathcal{F}(f)(\kappa) = \hat{f}(\kappa) = \int_{\mathbb{R}} e^{-2\pi i \kappa x} f(x) dx. \quad (A.1)$$

The operator \mathcal{F} can be extended to a unitary transformation on $L^2(\mathbb{R})$ with the Plancherel formula:

$$\int_{\mathbb{R}} \hat{f}(\kappa) \overline{\hat{g}(\kappa)} d\kappa = \int_{\mathbb{R}} f(x) \overline{g(x)} dx,$$

for any functions $f, g \in L^2(\mathbb{R} \rightarrow \mathbb{C})$.

We define the Hilbert transform of any function $f \in L^1(\mathbb{R})$,

$$\mathcal{H}(f)(x) = \frac{1}{\pi} \oint_{\mathbb{R}} \frac{f(u)}{x - u} du, \quad (A.2)$$

where the integral is defined in the principal value sense. The Hilbert transform can also be extended to a bounded operator on $L^2(\mathbb{R})$ which satisfies:

$$\widehat{\mathcal{H}(f)}(\kappa) = -i \operatorname{sgn}(\kappa) \hat{f}(\kappa) \quad (A.3)$$

where $\text{sgn}(\cdot)$ is the sign function. In particular, the identity (A.3) implies that \mathcal{H} is invertible on $L^2(\mathbb{R})$ and $\mathcal{H}^{-1} = -\mathcal{H}$. Moreover, let us mention that if $f \in L^1 \cap L^2(\mathbb{R})$ is absolutely continuous (i.e. $f' \in L^1(\mathbb{R})$), then the Hilbert transform of the function f is differentiable on \mathbb{R} and

$$\frac{d\mathcal{H}(f)}{d\kappa} = \mathcal{H}(f'). \quad (\text{A.4})$$

We define the Cauchy transform of any function $f \in L^1(\mathbb{R})$,

$$\mathcal{C}(f)(z) = \frac{1}{2\pi i} \int_{\mathbb{R}} \frac{f(x)}{x - z} dx. \quad (\text{A.5})$$

This function is analytic in both the lower and upper half planes, denoted \mathbb{C}_{\pm} . Moreover, its boundary values are given (in L^2 or pointwise if the limits make sense) by the Plemelj-Sokhotski formula, for all $x \in \mathbb{R}$,

$$\mathcal{C}(f)_{\pm}(x) = \pm \frac{f(x)}{2} + \frac{i}{2} \mathcal{H}(f)(x). \quad (\text{A.6})$$

We define the Sobolev space

$$H^{1/2}(\mathbb{R}) = \left\{ f \in L^2(\mathbb{R} \rightarrow \mathbb{R}) : \int_{\mathbb{R}} |\kappa| |\hat{f}(\kappa)|^2 d\kappa < \infty \right\}. \quad (\text{A.7})$$

This is a Hilbert space equipped with the inner-product

$$\langle f; g \rangle_{H^{1/2}} = \int_{\mathbb{R}} |\kappa| \hat{f}(\kappa) \hat{g}(-\kappa) d\kappa. \quad (\text{A.8})$$

There are other formulae for the inner product (A.8) which do not involve the Fourier transform. For any functions $f, g \in C^1(\mathbb{R})$ such that $f, g \in L^2(\mathbb{R})$, we claim that

$$\langle f; g \rangle_{H^{1/2}} = \frac{1}{4\pi^2} \iint_{\mathbb{R}^2} \frac{f(x) - f(y)}{x - y} \frac{g(x) - g(y)}{x - y} dx dy \quad (\text{A.9})$$

$$= \frac{-1}{2\pi} \int_{\mathbb{R}} f'(u) \mathcal{H}(f)(u) du. \quad (\text{A.10})$$

It can be checked that formula (A.9) holds for any functions $f, g \in H^{1/2}(\mathbb{R})$, while (A.10) holds as long as $\mathcal{F}(f') \in L^2(\mathbb{R})$ and $\mathcal{F}(f')(\kappa) = 2\pi i \kappa \hat{f}(\kappa)$.

We define Ξ to be the Gaussian noise associated with the Hilbert space $H^{1/2}(\mathbb{R})$, see for instance [39]. That is $\Xi = \{\Xi(f)\}_{f \in H^{1/2}(\mathbb{R})}$ is a Gaussian process with covariance structure:

$$\mathbb{E}[\Xi(f)\Xi(g)] = \langle f; g \rangle_{H^{1/2}}. \quad (\text{A.11})$$

Observe that, if $\chi_u(x) = \pi \mathbb{1}_{|x-u| \leq \ell/2}$, then

$$\widehat{\chi}_u(\kappa) = e^{-2\pi i u \kappa} \frac{\sin(\pi \ell \kappa)}{\kappa}. \quad (\text{A.12})$$

So that, if we define $\widehat{Q}(\kappa) = \frac{\sin^2(\pi \ell \kappa)}{|\kappa|}$, it is easy to check that despite the fact the functions $\chi_u \notin H^{1/2}(\mathbb{R})$, we have

$$\langle \chi_u; \chi_v \rangle_{H^{1/2}} = Q(u - v) \quad (\text{A.13})$$

for all $u \neq v$. Hence, in a *formal* sense, the log-correlated Gaussian field G with zero mean and covariance function Q can be realized as $G(u) = \Xi(\chi_u)$. This is only formal because the function χ_u does not belong to the domain of the noise Ξ . However we may rigorously define regularizations of the field G using this procedure. For any $\phi \in \mathcal{D}_0$ as in (1.30), and $\epsilon > 0$, we let

$$G_{\phi,\epsilon}(u) = \Xi(\chi_u * \phi_\epsilon). \quad (\text{A.14})$$

Then, by formulae (A.8), (A.12), and the definition of Q , the field $G_{\phi,\epsilon}$ has the correlation structure:

$$\begin{aligned} \mathbb{E}[G_{\phi,\epsilon}(u), G_{\psi,\delta}(v)] &= \langle \chi_u * \phi_\epsilon; \chi_v * \psi_\delta \rangle_{H^{1/2}} \\ &= \int_{\mathbb{R}} e^{-2\pi i(u-v)\kappa} \hat{\psi}(\delta\kappa) \overline{\hat{\phi}(\epsilon\kappa)} \widehat{Q}(\kappa) d\kappa \end{aligned} \quad (\text{A.15})$$

for any $\phi, \psi \in \mathcal{D}_0$, $\epsilon, \delta > 0$ and $u, v \in \mathbb{R}$. Note that by Plancherel's formula, we may also rewrite formula (A.15) as

$$\mathbb{E}[G_{\phi,\epsilon}(u), G_{\psi,\delta}(v)] = \iint_{\mathbb{R}^2} \phi_\epsilon(u-x) \psi_\delta(v-y) Q(x-y) dx dy.$$

Then, when it is convenient, we shall also denote the field $G_{\phi,\epsilon}$ by $G * \phi_\epsilon$. Moreover, the log-correlated field G can be realized as $G = \lim_{\epsilon \rightarrow 0} G_{\phi,\epsilon}$ in the sense of random distributions. It is not difficult to make this convergence rigorous, *e.g.* by using the asymptotics of Sect. 2.3 for the covariance kernel (A.15), but we do not pursue this here.

B. Gaussian Multiplicative Chaos

In this section, we review in further detail the theory of Gaussian Multiplicative Chaos (GMC) with respect to the Lebesgue measure on a compact subset $\mathfrak{A} \subset \mathbb{R}^d$. This theory which originates in the work of Mandelbrot and Kahane [41, 42, 49, 50] aims at defining the exponential of a log-correlated random field G , denoted formally by

$$\nu^\gamma(dx) = e^{\gamma G(x) - \frac{\gamma^2}{2} \mathbb{E}(G(x)^2)} dx. \quad (\text{B.1})$$

The original motivation to study such an object goes back to the work of Kolmogorov, who proposed that the measure ν^γ should describe the energy dissipation in a turbulent fluid; c.f. [58] for a modern reference. Another motivation comes from the fact that ν^γ can be interpreted as the Boltzmann–Gibbs measure associated with the *random Hamiltonian* G . Then, this measure describes the equilibrium configuration of a particle in a very rough landscape and $\gamma > 0$ plays the role of the inverse-temperature and is usually called the intermittency parameter. In fact, sampling from the measure ν^γ gives information about the points where the field G takes unusually high values known as γ -thick points, see [36]. In particular, there exists a critical value of γ above which no such points exist and the measure (B.1) needs to be renormalized in a different way and becomes purely atomic. This is known as the freezing transition in the theory of spin glasses and it has been observed that the behavior of ν^γ at criticality is related to the law of the maximum of the field G [25, 27]. Recently, there have also been intense developments in the case where G is the Gaussian Free Field associated to a domain in the complex plane. Then, the chaos measure (B.1) is known as the Liouville measure and it has been one of the key inputs in a program aiming at giving a mathematically

rigorous construction of Liouville quantum gravity and Liouville quantum field theory, c.f. the recent results of [16] and [51]. The latter paper is concerned with developing an important program of *imaginary geometries* and *Liouville quantum gravity and the Brownian map* which aims at proving that the Liouville measures are central objects which arise, for instance to describe the scaling limit of random planar maps. In addition various KPZ relations have been established [3, 22, 59]. For a more detailed introduction and further references to these topics, we refer to the lecture notes [4, 32, 61] or the survey [60]. Measures of the type (B.1) also started to play an increasingly important role in random matrix theory [6, 70] and in statistics of the Riemann zeta function high up on the critical line [64].

Usually, the random measure ν^γ is defined by first regularizing the field G in some way and then by taking a limit as the regularization tends to zero.

There have been many important developments in understanding this procedure and the so-called subcritical case is now well understood for Gaussian regularizations [5, 41, 62, 65]. To be specific, let G be a Gaussian process on \mathfrak{A} with a covariance kernel:

$$T(x, y) = -\log |x - y| + g(x, y)$$

where the function $g : \mathfrak{A}^2 \rightarrow \mathbb{R} \cup \{-\infty\}$ is continuous as an extended function and there exists a constant $C > 0$ so that for all $x, y \in \mathfrak{A}$,

$$g(x, y) - \log^+ |x - y| \leq C. \quad (\text{B.2})$$

We introduce this general setting since our main example is a stationary Gaussian process on \mathbb{R} with covariance kernel Q given by (1.22). Note that one usually assumes that g is a continuous and bounded function, but these assumptions can be relaxed without changing the general theory, as long as a condition such as (B.2) holds. Because of the singularity of the kernel T on the diagonal, G is not defined pointwise and needs to be interpreted as a random generalized function. Formally $G = \{G(f)\}_{f \in C(\mathfrak{A})}$ is a Gaussian process with covariance structure:

$$\mathbb{E}[G(f)G(g)] = \iint_{\mathbb{R}^2} f(x)g(y)T(x, y)dx dy.$$

In general, the definition of G can be extended to a more general class of test functions than $C(\mathfrak{A})$, or even to certain classes of measures. To define the exponential measure (B.1), one can consider a regularization of the process G coming from the convolution with an approximate delta function. This approach was introduced by Robert and Vargas in [62] and developed further by Berestycki in [5]. Namely, given $\epsilon > 0$ and a mollifier ϕ (i.e. a sufficiently smooth and light-tailed probability density function on \mathbb{R}), we define

$$G_{\phi, \epsilon} = G * \phi_\epsilon \quad \text{where} \quad \phi_\epsilon(x) = \epsilon^{-1}\phi(x/\epsilon),$$

and for any $\gamma > 0$,

$$\nu_\epsilon^\gamma(dx) = \exp\left(\gamma G_\epsilon(x) - \frac{\gamma^2}{2}\mathbb{E}\left[G_\epsilon(x)^2\right]\right)dx. \quad (\text{B.3})$$

Then, for any function $w \in L^1(\mathfrak{A} \rightarrow \mathbb{R}_+)$ uniformly bounded, if $\gamma^2 < 2d$, we let

$$\nu^\gamma(w) = \lim_{\epsilon \rightarrow 0} \nu_\epsilon^\gamma(w). \quad (\text{B.4})$$

The random measures ν^γ are called the multiplicative chaos measures associated to the Gaussian process with covariance kernel T . The major achievements of the GMC theory are that the limit (B.4) exists in probability (and almost surely in certain cases) and that it does not depend on ϕ for a large class of mollifiers. Moreover, the measure ν^γ is non-trivial if and only if $\gamma^2 < 2d$. In the critical ($\gamma = \sqrt{2d}$) and supercritical ($\gamma^2 > 2d$) regimes, one needs different normalizations than (B.3) to make sense of the GMC in a non-trivial way, c.f. [60, Section 6] and reference therein. In this paper, we will focus on the subcritical regime, in which case by Theorem 1.2, the limit (B.4) holds in $L^1(\mathbb{P})$ and the normalization is such that

$$\mathbb{E}[\nu^\gamma(w)] = \int w(u) du.$$

Moreover, for any $q \in \mathbb{N}$ such that $q\gamma^2 < 2$, it is not difficult to show that

$$\mathbb{E}[\nu^\gamma(w)^q] := \lim_{\epsilon \rightarrow 0} \mathbb{E}[\nu_\epsilon^\gamma(w)^q] = \int_{\mathfrak{A}^q} \exp\left(\gamma^2 \sum_{1 \leq j < k \leq q} T(u_j, u_k)\right) \prod_{k=1}^q w(u_k) du_k. \quad (\text{B.5})$$

In particular, in dimension $d = 1$, by a change of variables, this implies that for any $x_0 \in \mathbb{R}$,

$$\begin{aligned} & \mathbb{E}[\nu^\gamma([x_0 - \frac{r}{2}, x_0 + \frac{r}{2}])^q] \\ &= r^{\xi(q)} \int_{[0,1]^q} \prod_{1 \leq j < k \leq q} |u_j - u_k|^{-\gamma^2} e^{\gamma^2 g(x_0 + r(u_j - \frac{1}{2}), x_0 + r(u_k - \frac{1}{2}))} \prod_{k=1}^q du_k, \end{aligned}$$

where $\xi(q) = q - \gamma^2 \frac{q(q-1)}{2}$, so that

$$\mathbb{E}[\nu^\gamma([x_0 - \frac{r}{2}, x_0 + \frac{r}{2}])^q] \sim r^{\xi(q)} \mathcal{S}(q; \gamma^2/2) e^{\gamma^2 (\frac{q}{2}) g(x_0, x_0)} \quad \text{as } r \rightarrow 0,$$

where

$$\mathcal{S}(n; \tilde{\gamma}) := \int_{[0,1]^n} \prod_{i \neq j} |u_i - u_j|^{-\tilde{\gamma}} \prod_{k=1}^n du_k = \prod_{j=0}^{n-1} \frac{\Gamma(1 + j\tilde{\gamma})^2 \Gamma(1 + \tilde{\gamma} + j\tilde{\gamma})}{\Gamma(2 + (n + j - 1)\tilde{\gamma}) \Gamma(1 + \tilde{\gamma})} \quad (\text{B.6})$$

is a Selberg integral. The quadratic function $\xi(q)$ is called the structure exponent and it describes the multi-fractal properties of the random measure ν^γ , c.f. [60, section 2.3]. Finally, since the Selberg integral converges if and only if $\Re\{n\tilde{\gamma}\} < 1$, this shows that the condition $q\gamma^2 < 2$ in (B.5) is sharp.

Let us conclude this section by stating a result about the convergence of the GMC measure in the so-called L^2 -phase ($\gamma^2 < d$). In particular, we will need this result in Sect. 2.1 in order to identify the law of the multiplicative chaos measure coming from counting statistics of the CUE or sine process. In addition, the strategy of the proof will be re-used and generalized in Sects. 2.1 and 2.2 to construct multiplicative chaos measures for the asymptotically Gaussian fields discussed in the introduction.

Proposition B.1. *Let q be an even integer and $\gamma > 0$ so that $q\gamma^2 < 2d$. For any $w \in L^1(\mathfrak{A})$ uniformly bounded, the random variable $v_{Q,\epsilon}^\gamma(w)$ given by (B.3) converges as $\epsilon \rightarrow 0$ in $L^q(\mathbb{P})$ to a random variable $v_Q^\gamma(w)$ which does not depend on the mollifier ϕ subject to the conditions that ϕ is smooth and $\phi \in \mathcal{D}_\alpha$ for some sufficiently large $\alpha > 0$, (1.30).*

Proof. This follows from a standard argument; c.f. for instance the proof of [61, Theorem 2.3]. First, given a mollifier ϕ , one can prove that $v_\epsilon^\gamma(w)$ is a Cauchy sequence in $L^q(\mathbb{P})$. When q is an even integer, this just boils down to checking that

$$\mathbb{E}[G_{\phi,\epsilon}(x)G_{\phi,\epsilon'}(x')] \rightarrow T(x, x') \quad (\text{B.7})$$

as $\epsilon, \epsilon' \rightarrow 0$ in such a way that we can apply the dominated convergence theorem. To prove that the limit does not depend on ϕ , if $v_\epsilon^{\gamma'}$ is the measure associated to the Gaussian process $G_{\psi,\epsilon}$ for a second mollifier ψ , then it suffices to show that

$$\lim_{\epsilon \rightarrow 0} \mathbb{E}[|v_\epsilon^{\gamma'}(w) - v_\epsilon^\gamma(w)|^q] = 0.$$

This limit follows in a similar fashion by checking that as $\epsilon \rightarrow 0$,

$$\mathbb{E}[G_{\phi,\epsilon}(x)G_{\psi,\epsilon}(x')] \rightarrow T(x, x').$$

□

Remark B.2. In what follows, the condition (B.7) is replaced by Assumption 2.2. In fact, by going carefully through the proof in [5], it is not difficult to check that if

$$T_{\epsilon,\delta}(x, x') = \mathbb{E}[G_{\phi,\epsilon}(x)G_{\phi,\delta}(x')]$$

satisfies Assumption 2.2, then for any $\gamma < \sqrt{2d}$, $v_\epsilon^\gamma(w)$ converges in $L^1(\mathbb{P})$ to $v^\gamma(w)$ as $\epsilon \rightarrow 0$. Moreover, for the stationary Gaussian process on \mathbb{R} with covariance kernel Q given by (1.22) which arises in Theorem 1.3 for the mesoscopic CUE, as well as in Theorem 1.9 for the sine process, it is proved in Sect. 2.3 that Assumption 2.2 holds for any mollifier $\phi \in \mathcal{D}_\alpha$ for any $\alpha > 0$.

Acknowledgements. We thank Christian Webb for his interest in our work, for useful discussions and carefully reading the first draft of this paper which helped fix typos and mistakes in the proofs and greatly improved the presentation. G. L. thanks Juhan Aru for stimulating discussions about Gaussian multiplicative chaos. N. S. would like to thank Yan Fyodorov for suggesting to study the conjectures of [55] in the context of random matrix theory. Finally, we express our gratitude to the anonymous referees for their detailed comments and suggestions.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Arguin, L.-P., Belius, D., Bourgade, P.: Maximum of the characteristic polynomial of random unitary matrices. *Commun. Math. Phys.* **349**, 703–751 (2017)
2. Bekerman, F., Lodhia, A.: Mesoscopic Central Limit Theorem for General β -Ensembles. [arXiv:1605.05206](https://arxiv.org/abs/1605.05206) (2016)

3. Benjamini, I., Schramm, O.: KPZ in one dimensional random geometry of multiplicative cascades. *Commun. Math. Phys.* **289**(2), 653–662 (2009)
4. Berestycki, N.: Introduction to the Gaussian Free Field and Liouville Quantum Gravity. <http://www.statslab.cam.ac.uk/~beresty/Articles/oxford.pdf>. Accessed 18 Jan 2018 (2016)
5. Berestycki, N.: An elementary approach to Gaussian multiplicative chaos. *Electron. Commun. Probab.* **22**(27), 12 (2017)
6. Berestycki, N., Webb, C., Wong, M.D.: Random Hermitian matrices and Gaussian multiplicative Chaos. *Probab. Theory Relat. Fields* (2017)
7. Berggren, T., Duits, M.: Mesoscopic fluctuations for the thinned circular unitary ensemble. *Math. Phys. Anal. Geom.* (2017)
8. Borodin, A.: Determinantal point processes. In: *The Oxford Handbook of Random Matrix Theory*, pp. 231–249. Oxford University Press, Oxford (2011)
9. Borodin, A., Okounkov, A.: A Fredholm determinant formula for Toeplitz determinants. *Integral Equ. Oper. Theory* **37**(4), 386–396 (2000)
10. Bourgade, P., Erdős, L., Yau, H.-T., Yin, J.: Fixed energy universality for generalized Wigner matrices. *Commun. Pure Appl. Math.* **69**(10), 18151881 (2015)
11. Bourgade, P., Kuan, J.: Strong Szegő asymptotics and zeros of the zeta-function. *Commun. Pure Appl. Math.* **67**(6), 1028–1044 (2014)
12. Breuer, J., Duits, M.: Universality of mesoscopic fluctuations for orthogonal polynomial ensembles. *Commun. Math. Phys.* **342**(2), 491–531 (2016)
13. Chhaibi, R., Najnudel, J., Madaule, T.: On the Maximum of the $C\beta E$ Field. [arXiv:1607.00243](https://arxiv.org/abs/1607.00243) (2016)
14. Claeys, T., Krasovsky, I.: Toeplitz determinants with merging singularities. *Duke Math. J.* **164**(15), 2897–2987 (2015)
15. Cremers, H., Kadelka, D.: On weak convergence of integral functionals of stochastic processes with applications to processes taking paths in L^E_p . *Stoch. Process. Appl.* **21**(2), 305–317 (1986)
16. David, F., Kupiainen, A., Rhodes, R., Vargas, V.: Liouville quantum gravity on the Riemann sphere. *Commun. Math. Phys.* **342**(3), 869–907 (2016)
17. Deift, P.: Integrable operators. In: *Differential Operators and Spectral Theory*, Volume 189 of American Mathematical Society Translations: Series 2, pp. 69–84. American Mathematical Society, Providence (1999)
18. Deift, P., Its, A., Krasovsky, I.: Toeplitz matrices and Toeplitz determinants under the impetus of the Ising model: some history and some recent results. *Commun. Pure Appl. Math.* **66**(9), 1360–1438 (2013)
19. Deift, P., Kriecherbauer, T., McLaughlin, K.T.-R., Venakides, S., Zhou, X.: Uniform asymptotics for polynomials orthogonal with respect to varying exponential weights and applications to universality questions in random matrix theory. *Commun. Pure Appl. Math.* **52**(11), 1335–1425 (1999)
20. Deift, P., Zhou, X.: A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation. *Ann. Math. (2)* **137**(2), 295–368 (1993)
21. Ding, J., Roy, R., Zeitouni, O.: Convergence of the centered maximum of log-correlated Gaussian fields. *Ann. Probab.* **45**(6A), 3886–3928 (2017)
22. Duplantier, B., Sheffield, S.: Liouville quantum gravity and KPZ. *Invent Math.* **185**(2), 333–393 (2011)
23. Erdős, L., Knowles, A.: The Altshuler–Shklovskii formulas for random band matrices II: the general case. *Ann. Henri Poincaré* **16**(3), 709–799 (2015)
24. Forrester, P.J., Warnaar, S.O.: The importance of the Selberg integral. *Bull. Am. Math. Soc. (N.S.)* **45**(4), 489–534 (2008)
25. Fyodorov, Y.V., Bouchaud, J.-P.: Freezing and extreme-value statistics in a random energy model with logarithmically correlated potential. *J. Phys. A* **41**(37), 372001, 12 (2008)
26. Fyodorov, Y.V., Hiary, G.A., Keating, J.P.: Freezing transition, characteristic polynomials of random matrices, and the Riemann zeta function. *Phys. Rev. Lett.* **108**, 170601 (2012)
27. Fyodorov, Y.V., Keating, J.P.: Freezing transitions and extreme values: random matrix theory, and disordered landscapes. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **372**(2007), 20120503, 32 (2014)
28. Fyodorov, Y.V., Khoruzhenko, B.A., Simm, N.J.: Fractional Brownian motion with Hurst index $H = 0$ and the Gaussian unitary ensemble. *Ann. Probab.* **44**(4), 2980–3031 (2016)
29. Fyodorov, Y.V., Le Doussal, P.: Moments of the position of the maximum for GUE characteristic polynomials and for log-correlated Gaussian processes. *J. Stat. Phys.* **164**(1), 190–240 (With Appendix I by Alexei Borodin and Vadim Gorin) (2016)
30. Fyodorov, Y.V., Le Doussal, P., Rosso, A.: Statistical mechanics of logarithmic REM: duality, freezing and extreme value statistics of $1/f$ noises generated by Gaussian free fields. *J. Stat. Mech. Theory Exp.* **2009**(10), P10005, 32 (2009)
31. Fyodorov, Y.V., Simm, N.J.: On the distribution of the maximum value of the characteristic polynomial of GUE random matrices. *Nonlinearity* **29**, 2837–2855 (2016)
32. Garban, C.: Quantum gravity and the KPZ formula [after Duplantier-Sheffield]. *Astérisque* (352):Exp. No. 1052, ix, 315–354 (2013). Séminaire Bourbaki. Vol. 2011/2012. Exposés 1043–1058

33. Geronimo, J.S., Case, K.M.: Scattering theory and polynomials orthogonal on the unit circle. *J. Math. Phys.* **20**(2), 299–310 (1979)
34. He, Y., Knowles, A.: Mesoscopic eigenvalue statistics of Wigner matrices. *Ann. Appl. Probab.* **27**(3), 1510–1550 (2017)
35. Hough, J.B., Krishnapur, M., Peres, Y., Virág, B.: Determinantal processes and independence. *Probab. Surv.* **3**, 206–229 (2006)
36. Hu, X., Miller, J., Peres, Y.: Thick points of the Gaussian free field. *Ann. Probab.* **38**(2), 896–926 (2010)
37. Hughes, C.P., Keating, J.P., O’Connell, N.: On the characteristic polynomial of a random unitary matrix. *Commun. Math. Phys.* **220**(2), 429–451 (2001)
38. Its, A.R., Izergin, A.G., Korepin, V.E., Slavnov, N.A.: Differential equations for quantum correlation functions. In: *Proceedings of the Conference on Yang–Baxter Equations, Conformal Invariance and Integrability in Statistical Mechanics and Field Theory*, vol 4, pp. 1003–1037 (1990)
39. Janson, S.: *Gaussian Hilbert Spaces*, Volume 129 of Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge (1997)
40. Johansson, K.: Random matrices and determinantal processes. In: *Mathematical Statistical Physics*, pp. 1–55. Elsevier B. V., Amsterdam (2006)
41. Kahane, J-P: Sur le chaos multiplicatif. *Ann. Sci. Math. Québec* **9**(2), 105–150 (1985)
42. Kahane, J-P., Peyrière, J.: Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.* **22**(2), 131–145 (1976)
43. Kallenberg, O.: *Foundations of Modern Probability. Probability and Its Applications* (New York). Springer, New York (2002)
44. Krasovsky, I.: Aspects of Toeplitz determinants. In: *Random Walks, Boundaries and Spectra*, Volume 64 of Progress in Probability, pp. 305–324. Birkhäuser/Springer, Basel (2011)
45. Kuijlaars, A.: Riemann–Hilbert analysis for orthogonal polynomials. In: *Orthogonal Polynomials and Special Functions* (Leuven, 2002), Volume 1817 of Lecture Notes in Math., pp. 167–210. Springer, Berlin (2003)
46. Lambert, G.: Mesoscopic fluctuations for unitary invariant ensembles. *Electron. J. Probab.* **23** (2018), Paper no. 7. [arXiv:1510.03641](https://arxiv.org/abs/1510.03641) (2016)
47. Lodhia, A., Simm, N.J.: Mesoscopic linear statistics of Wigner matrices. [arXiv:1503.03533](https://arxiv.org/abs/1503.03533) (2015)
48. Macchi, O.: The coincidence approach to stochastic point processes. *Adv. Appl. Probab.* **7**, 83–122 (1975)
49. Mandelbrot, B.B.: Possible refinement of the log-normal hypothesis concerning the distribution of energy dissipation in intermittent turbulence. In: Rosenblatt, M., Van Atta, C. (eds.) *Statistical Models and Turbulence*, vol. 12, pp. 333–351. Springer, New York (1972)
50. Mandelbrot, B.B.: Limit lognormal multifractal measures. In: Gotsman, E.A. et al. (ed.) *Frontiers of Physics: Landau Memorial Conference*, pp. 309–340. Pergamon, New York (1990)
51. Miller, J., Sheffield, S.: Liouville Quantum Gravity and the Brownian Map III: The Conformal Structure is Determined. [arXiv:1608.05391](https://arxiv.org/abs/1608.05391) (2016)
52. Ostrovsky, D.: Mellin transform of the limit lognormal distribution. *Commun. Math. Phys.* **288**(1), 287–310 (2009)
53. Ostrovsky, D.: Selberg integral as a meromorphic function. *Int. Math. Res. Not. IMRN* **2013**(17), 3988–4028 (2013)
54. Ostrovsky, D.: On Barnes beta distributions and applications to the maximum distribution of the 2D Gaussian free field. *J. Stat. Phys.* **164**(6), 1292–1317 (2016)
55. Ostrovsky, D.: On Riemann zeroes, lognormal multiplicative chaos, and Selberg integral. *Nonlinearity* **29**(2), 426–464 (2016)
56. Ostrovsky, D.: On Barnes beta distributions, Selberg integral and Riemann xi. *Forum Math.* **28**(1), 1–23 (Published electronically in 2014) (2016)
57. Paquette, E., Zeitouni, O.: The maximum of the CUE field. *International Mathematics Research Notices* (2017)
58. Pereira, R.M., Garban, C., Chevillard, L.: A dissipative random velocity field for fully developed fluid turbulence. *J. Fluid Mech.* **794**, 369–408 (2016)
59. Rhodes, R., Vargas, V.: KPZ formula for log-infinitely divisible multifractal random measures. *ESAIM Probab. Stat.* **15**, 358–371 (2011)
60. Rhodes, R., Vargas, V.: Gaussian multiplicative chaos and applications: a review. *Probab. Surv.* **11**, 315–392 (2014)
61. Rhodes, R., Vargas, V.: Lecture notes on Gaussian multiplicative chaos and Liouville Quantum Gravity. [arXiv:1602.07323](https://arxiv.org/abs/1602.07323) (2016)
62. Robert, R., Vargas, V.: Gaussian multiplicative chaos revisited. *Ann. Probab.* **38**(2), 605–631 (2010)
63. Rodgers, B.: A central limit theorem for the zeroes of the zeta function. *Int. J. Number Theory* **10**(2), 483–511 (2014)
64. Saksman, E., Webb, C.: The Riemann Zeta Function and Gaussian Multiplicative Chaos: Statistics on the Critical Line. [arXiv:1609.00027](https://arxiv.org/abs/1609.00027) (2016)

- 65. Shamov, A.: On Gaussian multiplicative chaos. *J. Funct. Anal.* **270**(9), 3224–3261 (2016)
- 66. Simon, B.: Orthogonal Polynomials on the Unit Circle. Part 1, Volume 54 of American Mathematical Society Colloquium Publications. American Mathematical Society, Providence (2005). Classical theory
- 67. Simon, B.: Trace Ideals and Their Applications, Volume 120 of Mathematical Surveys and Monographs. American Mathematical Society, Providence, second edition (2005)
- 68. Soshnikov, A.: The central limit theorem for local linear statistics in classical compact groups and related combinatorial identities. *Ann. Probab.* **28**(3), 1353–1370 (2000)
- 69. Soshnikov, A.: Determinantal random point fields. *Uspekhi Mat. Nauk* **55**(5(335)), 107–160 (2000)
- 70. Webb, C.: The characteristic polynomial of a random unitary matrix and Gaussian multiplicative chaos—the L^2 -phase. *Electron. J. Probab.* **20**(104), 21 (2015)

Communicated by P. Deift